

Journal of
Geophysical
Research

VOLUME 66

APRIL 1961

NUMBER 4

THE SCIENTIFIC PUBLICATION
OF THE AMERICAN GEOPHYSICAL UNION

Journal of Geophysical Research

An International Scientific Publication

OFFICERS OF THE UNION

LLOYD V. BERKNER, *President*
THOMAS F. MALONE, *Vice President*
A. NELSON SAYRE, *General Secretary*
WALDO E. SMITH, *Executive Secretary*

OFFICERS OF THE SECTIONS

Geodesy

CHARLES PIERCE, *President*
FLOYD W. HOUGH, *Vice President*
BUFORD K. MEADE, *Secretary*

Seismology

LEONARD M. MURPHY, *President*
JAMES A. PEOPLES, JR., *Vice President*
BENJAMIN F. HOWELL, JR., *Secretary*

Meteorology

THOMAS F. MALONE, *President*
GORDON E. DUNN, *Vice President*
WOODROW C. JACOBS, *Secretary*

Geomagnetism and Aeronomy

L. R. ALDREDGE, *President*
C. T. ELVEY, *Vice President*
J. HUGH NELSON, *Secretary*

Oceanography

WALTER H. MUNK, *President*
DONALD W. PRITCHARD, *Vice President*
EUGENE C. LAFOND, *Secretary*

Volcanology, Geochemistry, and Petrology

ALFRED O. C. NIER, *President*
FRANCIS J. TURNER, *Vice President*
IRVING FRIEDMAN, *Secretary*

Hydrology

WALTER B. LANGBEIN, *President*
WILLIAM C. ACKERMANN, *Vice President*
CHARLES C. McDONALD, *Secretary*

Tectonophysics

PATRICK M. HURLEY, *President*
LOUIS B. SLICHTER, *Vice President*
H. RICHARD GAULT, *Secretary*

Since January 1959 (Vol. 64, No. 1) the *Journal of Geophysical Research* has been published monthly by the American Geophysical Union, the U. S. National Committee of the International Union of Geodesy and Geophysics organized under the National Academy of Sciences-National Research Council as the U. S. national adhering body. Publication of this journal is supported by the National Science Foundation and the Carnegie Institution of Washington. The new monthly combines the type of scientific material formerly published in the bi-monthly *Transactions, American Geophysical Union*, and the quarterly *Journal of Geophysical Research*. The *Transactions, American Geophysical Union* will continue as a quarterly publication for Union business and items of interest to members of the Union.

BOARD OF EDITORS

Editors: PHILIP H. ABELSON and J. A. PEOPLES, JR.

ASSOCIATE EDITORS

1959-1961

HENRI BADER	T. NAGATA
K. E. BULLEN	FRANK PRESS
CONRAD P. MOOK	A. NELSON SAYRE
WALTER H. MUNK	MERLE A. TUVE

JAMES A. VAN ALLEN

1960-1962

JULIUS BARTELS	L. A. MANNING
V. V. BELOUSSOV	TOR J. NORDENSON
E. G. BOWEN	E. N. PARKER
JOHN E. CHAPPELEAR	GEORGE P. RIGSBY
G. D. GARLAND	WALTER O. ROBERTS
GORDON J. F. MACDONALD	C. N. TOUART

JAMES R. WAIT

1961-1963

FRANKLIN I. BADGLEY	ROBERT O. REID
HENRY G. BOOKER	BRUNO ROSSI
JOSEPH W. CHAMBERLAIN	GEORGE H. SUTTON
HERBERT FRIEDMAN	DAVID K. TODD
MARK F. MEIER	VICTOR VACQUIER

ARTHUR H. WAYNICK

The Editors of the *Journal of Geophysical Research* welcome original scientific contributions on the physics of the earth and its environment.

Manuscripts should be submitted in triplicate to J. A. Peoples, Jr., Department of Geology, University of Kansas, Lawrence, Kansas. Authors' institutions, if in the United States or Canada, are requested to pay a publication charge of \$25 per page, which, if honored, entitles them to 100 free reprints.

Subscriptions to the *Journal of Geophysical Research* and *Transactions, AGU*, are included in membership dues.

Nonmember subscriptions, *Journal of Geophysical Research*...\$30 for back volume of 1959, \$42 for back volume of 1960, \$4 for this issue; \$20 for the calendar year 1961.

Nonmember subscriptions, *Transactions, AGU*....\$4 per calendar year, \$1.25 per copy.

Subscriptions, renewals, and orders for back numbers should be addressed to American Geophysical Union, 1515 Massachusetts Ave., Northwest, Washington 5, D. C. Suggestions to authors are available on request.

Advertising Representative: Howland and Howland, Inc., 230 Park Ave., New York 17, N. Y.

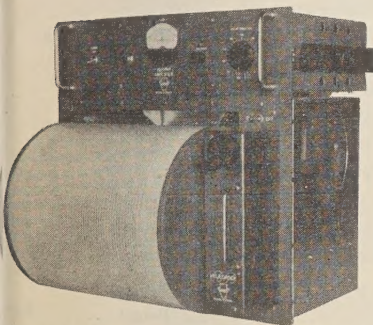
Published monthly by the American Geophysical Union from 1407 Sherwood Avenue, Richmond, Virginia. Second class postage paid at Richmond, Virginia.



PRECISION INSTRUMENTS for SEISMOLOGY — GEOPHYSICS

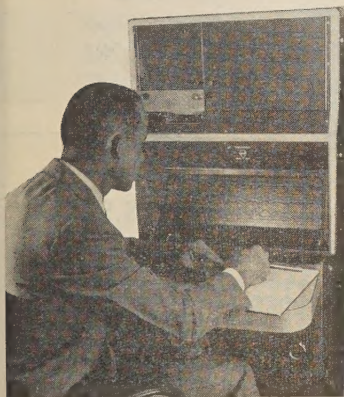
HELICORDER

An immediate data display, and up to 72 hours of data recorded on one 12" x 36" record are featured by this precision recorder. The penmotor(s) translates as the drum revolves, recording data on heat sensitive paper in helical fashion. Significant data is easily identified with one glance at the entire record. No dark room or delay for photo processing is necessary, and there is no ink to clog. One- and three-trace models are available. Drum speed is adjustable, and 8 to 72 hours of analog data in a band pass of 0 to 24 cps can be recorded on one permanent record. The companion amplifier shown is also available. The recorder can be rack-mounted or furnished with a cabinet for table mounting. Dependability is proven through years of trouble-free operation under extreme field conditions.



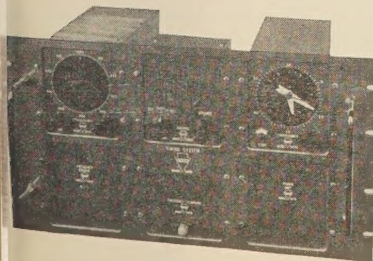
DEVELOCORDER

This is a combination 16 mm film recorder, automatic film processor, and film viewer. With 1 to 16 miniature Geotech galvanometers installed, Model 4000 will record up to 120 hours of data on a 200 ft. roll of film. Band pass is 0 to 120 cps depending upon film speed and galvanometers selected. With a cathode ray tube installed, Model 4000A will also record data across the film, as well as intensity-modulated spectrograms, Lissajous patterns, and other presentations. A film speed of 3 cm/min is standard, and speeds from 0.6 to 20 cm/min are available on special order. Self-contained chemicals process the film internally. Data is displayed magnified 10X on the 6" x 17" view-screen within 2 to 20 minutes after recording. A 2-directional, variable-speed drive is provided for viewing. Date and time are printed on the film at regular intervals.



PORTABLE STATION TIMING SYSTEM

This is a complete station timing system for precision recording and timing applications which has been transistorized and packaged in miniature modules for dependable, portable field use. It provides accurate clock time and programmed time marks, and 15 va of 60-cps power output, all with a stability of at least 5 parts in 10^7 per week. Individual modules, which can be obtained separately, are: an oven- and crystal-controlled frequency standard, a frequency divider, a strobe unit for comparing and adjusting system time with external time, a clock and time-mark programmer, a speaker for radio time signals, and an amplifier providing the output power for recorder drives and other devices. Dimensions are 19"W x 10.5"H x 9"D. Weight is 25 pounds. Power required is 2 amps, 24 volts, DC. These systems have operated in the field for over a year.



For detailed specifications, price, or delivery, write to:

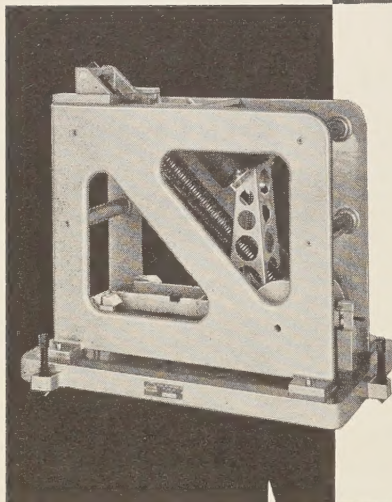
THE GEOTECHNICAL CORP.

401 SHILOH ROAD

GARLAND, TEXAS

Please mention JOURNAL OF GEOPHYSICAL RESEARCH, when writing to advertisers

SPRENGNETHER LONG PERIOD VERTICAL SEISMOMETER



WRITE FOR DETAILED
INFORMATION ON THESE
INSTRUMENTS.

GENERAL SPECIFICATIONS:

- Period Range: 6 to 70 seconds.
- Magnification: Up to 15,000, depending on operating period.
- Damping: Electromagnetic.
- Transducer-moving coils in circular magnetic gaps.
- Boom centering adjustment visible through sealed windows to facilitate adjustment.
- Coils approximately 500 ohms. or to your specifications.
- Pendulum steady mass weight 22 lbs.
- Invar spring for thermal stability.
- Airtight metal cover is provided to prevent recording of microbarometric oscillations.

PHYSICAL SPECIFICATIONS:

Length.....26" Width.....14" Height.....22"
Net Weight.....120 lbs. Shipping Weight.....175 lbs.

*To compliment this instrument,
a long period horizontal seismometer is also available.*

Internationally Known Mfrs. of Seismological, Geophysical Instruments.

W. F. SPRENGNETHER INSTRUMENT CO., INC.

4567 SWAN AVENUE

ST. LOUIS 10, MO.

BULLETIN (IZVESTIYA), ACADEMY OF SCIENCES, U.S.S.R. GEOPHYSICS SERIES

Subscriptions for 1960 series now available

This monthly Russian publication, perhaps the leading journal of Geophysics of the U.S.S.R., is being translated and published in an English edition for the year 1960 by the American Geophysical Union. The twelve numbers in Russian cover about 2000 pages. Published with the aid of a grant from the National Science Foundation.

Send subscriptions now to

AMERICAN GEOPHYSICAL UNION

1515 Massachusetts Avenue, N.W.

Washington 5, D. C., U.S.A.

Subscription rates: \$25.00 for the volume of 12 numbers (\$20.00 to individual members of AGU subscribing for personal use)

Numbers will be mailed as issued.

The English edition of this publication for 1957 has been translated and published for the American Geophysical Union by Pergamon Press. This volume may be ordered through the American Geophysical Union at a price of \$25.00. The 1958 and 1959 series are available at a price of \$25.00 for each volume of 12 numbers. Titles and authors of the papers contained in the series have been published in the 1959 and 1960 issues of the *Transactions*, AGU.

Please mention JOURNAL OF GEOPHYSICAL RESEARCH, when writing to advertisers



Gaspard de Coriolis: "A particle which is subject to no forces in a rotating coordinate system experiences a radial acceleration and a tangential acceleration."

It was around 1840 that Coriolis discovered what has since become known as the Coriolis Effect. He noticed objects above the earth tend to rotate relative to the earth's rotation . . . to the right in the northern hemisphere, to the left in the southern. The Coriolis Effect is in force in outer space, too. If a space vehicle is rotated in order to establish artificial gravity, the necessarily short radius of the rotation causes a Coriolis force. This creates orientation problems for a human occupant. To eliminate this difficulty, a scientist at Lockheed Missiles and Space Division conceived the idea of connecting the vehicle to an auxiliary fuel tank by a half-mile-long cable. Thus, if the whole system is then rotated at a reduced speed around its center of mass gravity, the longer radius greatly minimizes the Coriolis force. Right now—on the drawing boards at Lockheed—is an enormously advanced space vehicle system which utilizes this concept, in addition to many others.

Fortunately, natural laws are about the only restrictions which circumscribe scientists and engineers at Lockheed Missiles and Space Division. The climate in Sunnyvale and Palo Alto, on the San Francisco Peninsula, is close to perfection. The atmosphere—the opportunity to work on such important projects as the DISCOVERER and MIDAS satellites, the POLARIS FBM, or even more advanced concepts such as the space system cited above—is the dream of the creative engineer. Why not investigate future possibilities at Lockheed? Write Research and Development Staff, Dept. M-11G, 962 West El Camino Real, Sunnyvale, Calif. U.S. citizenship or existing Department of Defense industrial security clearance required.

Lockheed / **MISSILES AND SPACE DIVISION**

Programs Manager for the Navy POLARIS FBM and the Air Force AGENA Satellite in the DISCOVERER and MIDAS Programs

SUNNYVALE, PALO ALTO, VAN NUYS, SANTA CRUZ, SANTA MARIA, CALIFORNIA • CAPE CANAVERAL, FLORIDA • HAWAII

GEOPHYSICAL MONOGRAPH SERIES

AMERICAN GEOPHYSICAL UNION

1515 MASSACHUSETTS AVENUE, N.W.

WASHINGTON 5, D. C., U.S.A.

Antarctica in the International Geophysical Year—Geophysical Monograph No. 1 (Publication No. 462, National Academy of Sciences—National Research Council); Library of Congress Catalogue Card No. 56-60071; 133 pp. and large folded map of the Antarctic, 1956, 7" x 10", \$6.00. Contains 16 pages by various American authorities on the Antarctic under the headings: General, Geographic and Meteorological, Geological and Structural, Upper Atmospheric Physics, and Flora and Fauna. Map (41" x 41") compiled by the American Geographical Society. Introduction by L. M. Gould.

Geophysics and the IGY—Geophysical Monograph No. 2 (Publication No. 590, National Academy of Sciences—National Research Council); Library of Congress Catalogue Card No. 58-60035; 210 pp., 1958, 7" x 10", \$8.00. Contains 30 papers by leading American authorities under the headings: Upper Atmospheric Physics, The Lower Atmosphere and the Earth, and The Polar Regions. Preface by Joseph Kaplan.

Atmospheric Chemistry of Chlorine and Sulfur Compounds—Geophysical Monograph No. 3 (Publication No. 652, National Academy of Sciences—National Research Council); Library of Congress Catalogue Card No. 59-60039; 129 pp., 1959, 7" x 10", \$5.50. Based on a symposium held jointly with the Robert A. Taft Sanitary Engineering Center of the U. S. Public Health Service in Cincinnati in November, 1957. Contains 23 papers (some as summaries) with discussion. Preface by James P. Lodge, Jr.

Contemporary Geodesy—Geophysical Monograph No. 4 (Publication No. 708, National Academy of Sciences—National Research Council); Library of Congress Catalogue Card No. 59-60065; 96 pp., 7" x 10", 1959, \$5.50. Based on a Conference held at Cambridge, Massachusetts, in December 1958 jointly by the AGU with the Smithsonian Astrophysical Observatory and the Harvard College Observatory. Contains 14 papers by leading authorities, with verbatim discussions on topics ranging from classical geodesy to trilateration by underwater sound to space navigation in the solar system. Edited by Charles A. Whitten and Kenneth H. Drummond.

Physics of Precipitation—Geophysical Monograph No. 5 (Publication No. 746, National Academy of Sciences—National Research Council); Library of Congress Catalogue Card No. 60-60010; 435 pp., 7" x 10", 1960, \$12.50. Based on a Conference held at Woods Hole, Massachusetts, in June 1959. Contains 48 papers by leading authorities, with verbatim discussions on topics ranging from planetary-scale phenomena to microanalysis including hail formation and precipitation control. Edited by Helmut Weickmann.

Postage is to be added to prices shown unless payment accompanies order. Quantity discounts (count each Monograph separately): 5-19 copies, 10%; 20-49 copies, 15%; 50 or more copies 20%.

Purchase Order

TO AMERICAN GEOPHYSICAL UNION

1515 Massachusetts Avenue, N.W., Washington 5, D. C., U.S.A.

Please enter our order for the following:

_____ copies of Geophysical Monograph No. 1, at \$6.00 *	\$ _____
_____ copies of Geophysical Monograph No. 2, at \$8.00 *	\$ _____
_____ copies of Geophysical Monograph No. 3, at \$5.50 *	\$ _____
_____ copies of Geophysical Monograph No. 4, at \$5.50 *	\$ _____
_____ copies of Geophysical Monograph No. 5, at \$12.50 *	\$ _____

☐ Payment of \$ _____ is enclosed.

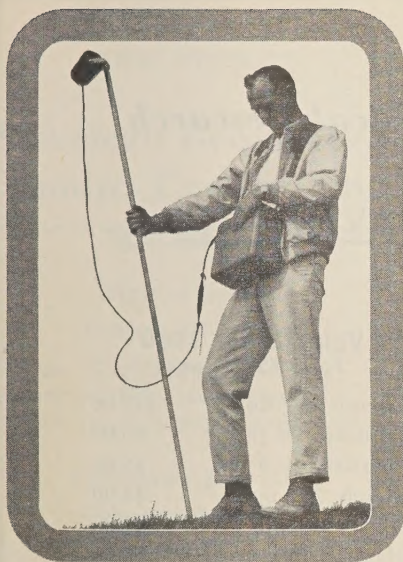
☐ Please send invoice, adding postage charges.

☐ Enter our standing order for _____ copies of subsequent Geophysical Monographs at the special prepublication rates, e.g., prepublication rate for Monograph No. 4 for non-members was \$4.00, payment in advance, or \$4.75 (plus postage) on invoice.

* List price is net for quantities up to four; see above for discounts on quantity purchases. Special discounts to members.

Typed name _____ Signature _____

Address _____



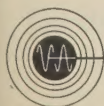
**M-49A PORTABLE
MAGNETOMETER
AVAILABLE NOW AT
\$3,950**

**PROVIDES FAST, ACCURATE
MAGNETIC SURVEYS
AT STILL LOWER COST**

World-wide acceptance of Russell Varian's *proton free precession principle* and wide usage of the M-49 portable magnetometer have now made it possible for Varian Associates to price the improved M-49A within the budget of most exploration programs.

PORTABLE The Varian M-49A, a complete proton magnetometer weighing only 22 pounds, furnishes direct readings in gammas every six seconds. Sensitive to better than ± 10 gammas, it requires no calibration or levelling and is so versatile it can make equally accurate field surveys on land, in the air, or over water — throughout the world.

Immediate delivery, parts, and service available in many countries of the free world. Technical data, sales, and lease information from Instrument Division:



VARIAN associates
PALO ALTO 40, CALIFORNIA

HYDROMECHANICISTS

Hydraulic Engineers

COASTAL ENGINEERS

*Conduct Experimental and
Theoretical Research with
the U. S. Navy into Ocean
Forces and Pressures.*

If you are interested in studying the reactions to water pressures of such obstructions as moored platforms, piers, breakwaters and harbors, contact the Commanding Officer and Director.

U. S. NAVAL Civil Engineering LABORATORY

Department J

Port Hueneme, California

(On the coast between Los Angeles and
Santa Barbara)

Starting Salary up to \$10,635 with regular increases, plus benefits worth up to \$2,563 per year. Career Civil Service.

Journal of Geophysical Research

BACK ISSUES AVAILABLE

Volume 64, 1959

Total 2488 pages

Complete Volume	\$30.00
January, 132 pp.	\$2.00
February, 138 pp.	\$2.00
March, 112 pp.	\$2.00
April, 106 pp.	\$2.00
May, 98 pp.	\$2.00
June, 110 pp.	\$2.00
July, 168 pp.	\$2.00
August, 268 pp.	\$4.00
September, 230 pp.	\$3.00
October, 284 pp.	\$4.00
November, 390 pp.	\$5.00
December, 452 pp.	\$6.00

Volume 65, 1960

Total 4248 pages

Complete Volume	\$42.00
January, 384 pp.	\$5.00
February, 414 pp.	\$5.00
March, 284 pp.	\$4.00
April, 248 pp.	\$4.00
May, 314 pp.	\$4.00
June, 220 pp.	\$4.00
July, 348 pp.	\$4.00
August, 350 pp.	\$4.00
September, 462 pp.	\$6.00
October, 490 pp.	\$6.00
November, 344 pp.	\$4.00
December, 390 pp.	\$5.00

SYMPOSIA REPRINTED from JGR

International Symposium on Electronic Distance-Measuring Techniques (144-page Symposium reprinted from the February 1960 issue)	\$3.50
Symposium on Sferics and Thunderstorm Electricity (102-page Symposium reprinted from the July 1960 issue)	\$3.50
Symposium on the Exosphere and Upper <i>F</i> Region (74-page Symposium reprinted from the September 1960 issue)	\$2.50
Scientific Effects of Artificially Introduced Radiations at High Altitudes (74-page Symposium reprinted from the August 1959 issue)	\$1.50
International Symposium on Fluid Mechanics in the Ionosphere (202-page Symposium reprinted from the December 1959 issue)	\$4.50

AMERICAN GEOPHYSICAL UNION

1515 Massachusetts Avenue, N.W., Washington 5, D. C.

To obtain back issues for 1958 and earlier years, write to
Walter J. Johnson, Inc., 111 Fifth Avenue, New York 3, New York.

NEW REPRINT

American Geophysical Union: Transactions

(Reproduced with the permission of the American
Geophysical Union)

Now Available

Volumes 13-15, 1932-1934

Volume 13, 1932, paper bound

Volume 14, 1933, paper bound

Volume 15, 1934, paper bound

Previously Reprinted

Volumes 1-12, 1920-1931

(Volumes 3 and 5 were never published)

Paper bound set (in 9 volumes) \$110.00

Volume 1, 1920, paper bound 5.00

Volume 2, 1921, paper bound 10.00

Volume 4, 1923, paper bound 15.00

Volume 6, 1925, paper bound 5.00

Volume 7, 1926, paper bound 15.00

Volume 8, 1927, paper bound 20.00

Volume 9, 1928, paper bound 15.00

Volume 10-11 1929-1930, paper
bound 20.00

Volume 12, 1931, paper bound 15.00

(Volumes 2, 4, and 6-9 published in
National Research Council Bulletin)

Volumes 16-34, 1935-1953, will be reproduced
by photo-offset as soon as there is sufficient
demand to warrant the understanding of a
reprint edition.



JOHNSON

REPRINT CORPORATION

111 FIFTH AVENUE

NEW YORK 3, NEW YORK

NASA—GODDARD
SPACE FLIGHT CENTER

experimental physicists and engineers

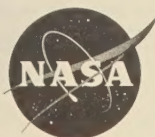
The Planetary Atmospheres Laboratory of the Goddard Space Flight Center offers stimulating and professionally rewarding positions for versatile experimental physicists and engineers. Duties include planning and execution of rocket and satellite experiments to measure atmospheric pressures, densities, temperatures, winds, and composition, including neutral particles, ions, and free-radicals.

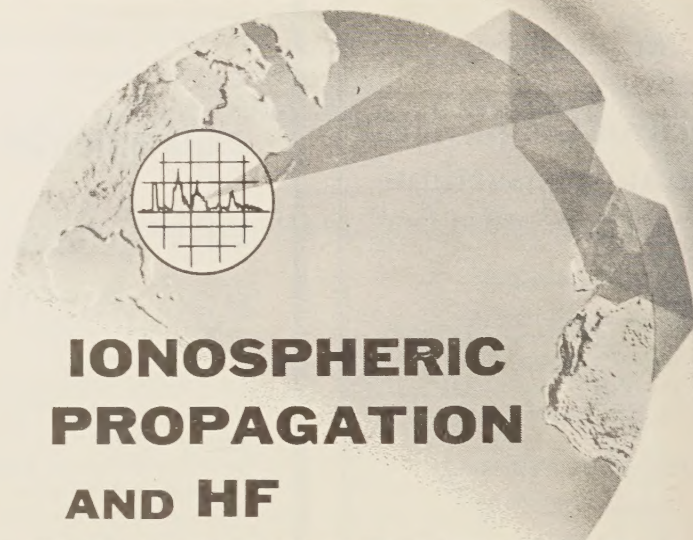
Results of these measurements will be used to describe the physics of the upper atmosphere. Appropriate general problems in physics, electronics, mechanics, and aerodynamics are involved; examples of specific topics are vacuum physics, neutral particle and ion mass spectrometry, light scattering, and molecular beam phenomena.

These positions require a Ph.D. degree in physics or engineering, or a Masters degree and suitable experience. For additional information, address your inquiry to:

N. W. Spencer,
Head, Planetary Atmospheres
NASA Goddard Space Flight Center
Greenbelt, Maryland
(Suburb of Washington, D.C.)

**National
Aeronautics
and Space
Administration**





IONOSPHERIC PROPAGATION AND HF COMMUNICATIONS

If you have the background, the imagination and the desire to contribute to important programs in these fields, you are invited to join a carefully selected team of outstanding scientists and engineers now contributing significantly to current knowledge through advanced research.

Our present needs are for:

SENIOR IONOSPHERIC PHYSICISTS

Ph.D. preferred, with several years' experience in the study of ionospheric phenomena. Should be familiar with present knowledge of upper atmosphere physics and possess an understanding of current programs using rockets and satellites for studies in F-region and beyond. Qualified individuals with supervisory abilities will have an exceptional opportunity to assume project leadership duties on HF projects already under way involving F-layer propagation studies backed by a substantial experimental program.

SENIOR DEVELOPMENT PHYSICISTS

Advanced degree in Physics or E.E. preferred. Must be familiar with latest techniques in the design of advanced HF receivers and transmitters and possess working knowledge of modern HF networks employing ferrites and metallic tape cores. Strong theoretical background in modern linear circuit theory desired. Will carry out laboratory development and implementation of new HF communications systems.

SENIOR ELECTRONIC ENGINEERS

Advanced degree in E.E. preferred. Must be familiar with conventional pulse circuit designs and applications. Technical background should include substantial experience in data process and data recovery systems using both analog and digital techniques. Knowledge of principles and application of modern information theory including correlation techniques helpful. Will be responsible for the design of sub-systems.

These programs are being conducted at our ELECTRO-PHYSICS LABORATORIES in the suburban Washington, D.C. area, ideally located from the viewpoint of advanced study which may be conducted at one of several nearby universities; for readily available housing in pleasant residential neighborhoods; and for the general amenities of living offered by this important Metropolitan center.

*For a prompt reply to your inquiry,
please forward resume in confidence to:*

W. T. WHELAN
Director of Research
& Development

ACF ELECTRONICS DIVISION
ACF INDUSTRIES
RIVERDALE, MARYLAND

Journal of GEOPHYSICAL RESEARCH

VOLUME 66

APRIL 1961

No. 4

The Time Variations of Solar Cosmic Rays during July 1959 at Minneapolis

J. R. WINCKLER, P. D. BHAVSAR, AND L. PETERSON

*School of Physics, University of Minnesota
Minneapolis 14, Minnesota*

Abstract. Thirteen high-altitude balloon flights were made at Minneapolis, Minnesota, during the period July 10–18, 1959. This paper summarizes measurements with ion chambers, Geiger counters, and scintillation counters of the solar cosmic rays accompanying three large flares during that period. Very large fluxes were observed beginning with the main phase of each of the three geomagnetic storms following the flares. During this main phase period, the instrumental results are consistent with protons having a steep energy spectrum and a lower limit at the air cutoff. For example, at 1046 UT on July 15 the integral energy spectrum for protons valid between 88 and 300 Mev was found to be

$$N(>E) = 1.05 \times 10^8 E^{-2.9} / \text{cm}^2 \cdot \text{sec} \cdot \text{ster}$$

The large increase observed in each storm can be interpreted as a rapid decrease in the cosmic-ray geomagnetic cutoff coincident with the start of the main phase of the storm. The flux returned to normal much sooner than the recovery of the main field of the field of the earth. At the time of the sudden commencement of the geomagnetic storm, a rapid and appreciable decrease of residual low-intensity solar cosmic rays was observed in two cases. It is suggested that these effects may be produced by ring current systems in the outer field of the earth associated with the geomagnetic storm. Following the July 15 increase, intensity oscillations in the solar proton stream with a period of $\frac{1}{2}$ hour were observed. These oscillations are of unknown origin.

Introduction. In this paper we should like to report observations made with scintillation and Geiger counters and ionization chambers carried by a series of balloon flights at altitudes of approximately 30 km during the three major solar cosmic-ray accelerations in July 1959. Because of the intermediate geomagnetic latitude of Minneapolis (where the flights were made), the solar cosmic-ray intensity is sensitively dependent upon the condition of the geomagnetic field, as has been demonstrated by a number of past events [Winckler, 1960; Winckler and Bhavsar, 1960]. The principal purpose of this paper is to describe the time variations of the solar cosmic rays as observed at Minneapolis and to associate these time variations with the behavior of the geomagnetic field.

The cosmic-ray particles directly followed

the appearance of three large solar flares occurring on July 10 at 0210 UT, on July 14 at 0325 UT, and on July 16 at 2114 UT [Reid and Leinbach, 1959]. At Minneapolis the cosmic-ray events were well separated. Each flare was followed, in 1 to 2 days, by a strong geomagnetic disturbance, during which the main intensity of the cosmic rays appeared. The solar cosmic-ray intensity at Minneapolis dropped to zero at the termination of each period of geomagnetic activity, and cosmic-ray levels remained at the galactic background until the next event began. However, precursor effects and many other fluctuations were observed during the storms. These will be discussed in detail for each of the three cases below. The time variations observed at Minneapolis are *not* characteristic of the polar-cap time variations. In the

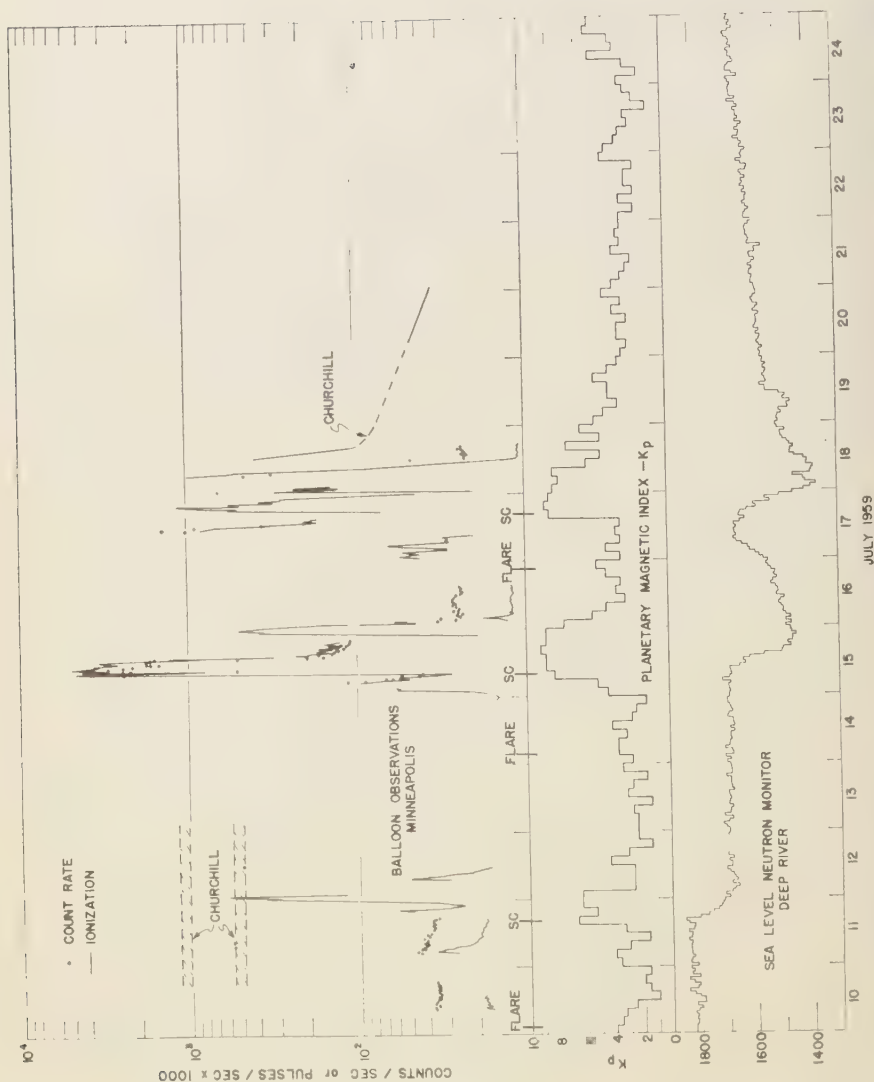


Fig. 1a. Upper: High-altitude total ionization (solid curve) and counter rate (dotted curve) during the three July events. Note logarithmic scale. Center: Note general correlation between influx of radiation and planetary K indices. Lower: Sea-level neutron intensity showing the three large Forbush decreases strongly correlated with the magnetic activity. On July 17 the flare cosmic rays produced a sea-level effect before the geomagnetic storm extending a range of at least 1 day.

polar regions prompt effects were noticed at the time of the flares; continuous large intensities that decayed slowly for many days and were unrelated to geomagnetic activity were also noticed.

The instrumentation used for these flights was mainly that developed during the IGY period for the Minnesota high-altitude monitoring program and has been described in detail in the literature [Winckler, 1960]. The data given in this paper are the ionization rate (referred to our standardized system [Winckler, 1960], the normalized Geiger counter rate, and, on many of the flights, scintillation counter data as described previously [Winckler and Bhavsar, 1960]. Some flights carried one or two of the instruments and some flights all three. All these flights and various others carried nuclear emulsions (Ney, Freier, and others, to be published).

Summary of the July events. Figure 1a summarizes the observations during the period. The various events will be discussed subsequently with detailed curves. Figure 1 plots in the upper section the Geiger counter rate and the total ionization rate taken from the high-altitude level flight portion of the balloon flights at Minneapolis. Also shown are the times of the main flares and sudden commencements of the magnetic storms; below this, the planetary magnetic indices on the 3-hour basis; and at the bottom, the sea-level cosmic-ray intensity given by the Deep River Neutron Monitor [Carmichael and Steljes, 1959]. The first flare in the series occurred early on July 10, and a balloon flight at high altitude shortly thereafter showed only normal rates during that day. The next flight, early on July 11, showed a transitory increase that died away, and no effect was observed at the time of the sudden commencement. A few hours after the sudden commencement, however, a large increase of approximately 50 times normal occurred. This died away and the high-altitude cosmic-ray rates returned toward normal. No flights were made for a period of 2 days until the next flare occurred early on July 14. A balloon flight made at the end of July 14 showed a moderate increase before the sudden commencement. This increase was on the decline when the sudden commencement occurred. One-half hour after the sudden commencement, the rate increased by more than two orders of magnitude above normal cosmic-

ray intensities and persisted for about 1 day. The solar cosmic rays then disappeared, and at high altitude the galactic cosmic-ray background rate again was observed to be lowered by the second Forbush-type event that had occurred. The next flare occurred late on July 16, and balloon flights shortly thereafter showed some rather strong transitory increases ahead of the magnetic storm. These increases were also observed on sea-level monitors [Carmichael and Steljes, 1959] (see also lower line of Fig. 1 on July 17), and imply a high-energy component in this event. Following the sudden commencement, a large series of increases was observed which died away by about the middle of July 18, at which time the cosmic-ray intensity for the galactic component registered a third Forbush decrease and was the lowest recorded throughout the period of sunspot maximum. On our normalized scale the ionization rates reached a value of 11.2. On the same scale the intensity at solar minimum was 35 [Winckler, 1960].

In Figure 1a it is clear that the large transitory increases of cosmic-ray intensity at Minneapolis correspond well to the increased values of the planetary magnetic indices characteristic of the stormy periods. At the same time, in the bottom line, we see that the galactic cosmic rays decrease at the time of the magnetic storm, which is the familiar Forbush phenomenon. In this case, three large Forbush decreases accompanying the three storms followed one another in rapid succession so that the recovery from one was not complete before the next one occurred. Furthermore, on July 17 there is evidence for the sea-level increase in the cosmic rays associated with the flare late on July 16. In this summary it is clear that the cosmic-ray increases are associated with the magnetic storms rather than with the flares themselves. It must be emphasized that this correlation is mainly a result of the geomagnetic effects of the solar plasma cloud operating on the cosmic-ray cutoff energies at the latitude of Minneapolis where the flights were made, and that it is *not* a property of the solar cosmic rays in space. That this hypothesis is correct can be proved only from the high-latitude observations made during this period. Although we shall not discuss these observations in this paper, there is ample evidence that solar cosmic rays traveled from the sun and were detected over the polar regions of the

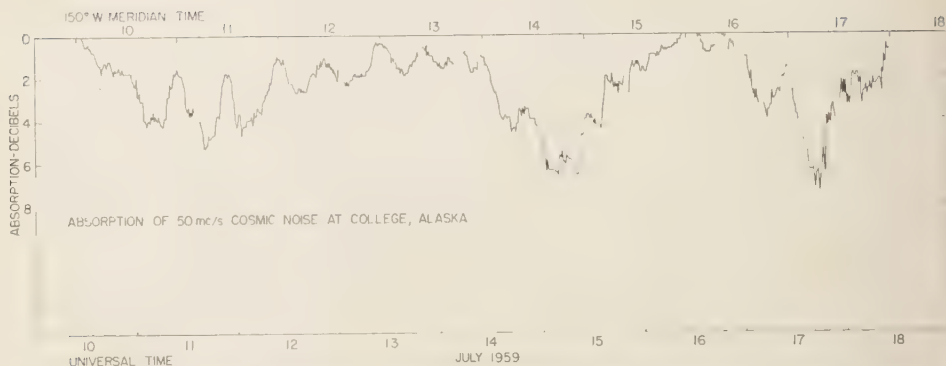


Fig. 1b. Polar-cap ionospheric absorption obtained from riometer at College, Alaska, courtesy of H. Leinbach. Note that the intensity grows and decays for each event in a smooth manner without correlation with the magnetic field. The semidiurnal fluctuations on the curve are the effect of sunlight on the ionosphere.

earth in the normally allowed range of latitudes, according to Störmer theory, within a few hours of each of these three flares, and that the intensity thereafter died away but lasted for many days following each flare event, so that there was overlap between events seen at high latitudes. This information, for example, comes from the ionospheric measurements of Reid and Leinbach [1959] (see Fig. 1b) and of Dana Bailey (private communication). Some high-latitude balloon measurements that are in agreement with early and continued arrival of these solar cosmic rays also were made, for example the results of Earl from balloon flights at Churchill (July 11 and 18, unpublished), the results of Anderson from balloon flights made at Resolute Bay on July 18 and thereafter [Anderson and Enemark, 1960], the results of Soviet measurements at Murmansk at high latitude [Charakhchian, Tulinov and Charakhchian, 1960], and measurements made at Fairbanks, Alaska, by Brown [Brown and D'Arcy, 1959]. We shall not discuss the very important question of the comparison of all these measurements at various geographical locations, but will confine ourselves for the present to documenting the results at Minneapolis.

Although the cosmic-ray events seem quite certainly associated with the three great flares of class 3+ shown in Figure 1a, this was a time of very high solar activity with several highly active regions on the solar disk generating many flares of classes 1, 2, and 2+. The region containing the great cosmic-ray flares

was McMath no. 5265 which appeared on the visible disk and became active about July 8 [National Bureau of Standards, 1959].

Table 1 summarizes the balloon flights made for the three flare events, and gives the instruments carried on each flight, the time of launch, and the duration of the constant level portion of the flights. All flights were launched from Minneapolis and drifted west, but all remained within about 1° of geomagnetic latitude for the duration of the flight.

In the detailed plots that follow, the rates of the instruments have been plotted as a function of time beginning at the time of launch of the balloon and continuing, if possible, until the parachute reaches the ground at the end of the flight. Accordingly, each of the curves shows an increasing rate which is associated with the ascent of the balloon through the atmosphere to its floating altitude at approximately 30 km or about 10-g/cm^2 depth below the top of the atmosphere. The part of the flight occupied by the ascent and the part occupied by the constant level performance are indicated. The constant level atmospheric depth is also indicated on all of these curves. The normal behavior for galactic cosmic radiation, for example, or flight IGC-12 on July 10, shows transition maxima located near 80-g/cm^2 depth in both the ionization chamber and counter due to the degradation of energy of the primary galactic cosmic rays in the atmosphere. In addition, low-energy nucleons are present, or other radiation which is absorbed in the high layers, marked

TABLE 1. Summary of Balloon Flights Made at Minneapolis, Minnesota, during July 1959 Solar Proton Events

Flight No.	Date UT	Balloon Launch UT	Duration of Constant Level UT	Atmospheric Depth g/cm ²	Instruments
IGC-12	July 10	0700	0905 to 1925	10.5	I.C.(A), G.C., Scint., Eml.
IGC-13	July 11	0252	0518 to 0700	10	I.C.(A), G.C., Scint., Eml.
			1250 to 1923	11	
M-A	July 11	1723	1952 to 2010	7	I.C.(B), Eml.
			2300 to 0150	10	
M-2	July 12	0503	0722 to 1120	6	I.C.(B), G.C., Eml.
IGC-F	July 14	2130	2320 to 2330	18	I.C.(A), Eml.
			0310 to 0315	28	
IGC-14	July 15	0325	0523 to 1700	9	I.C.(A), G.C., Scint., Eml.
IGC-G	July 15	0804	1046 to . . .	7	I.C.(C), G.C., Eml.
M-B	July 15	1807	2325 to 0310	28	I.C.(B), Eml.
IGC-15	July 16	0148	0402 to 0500	7	I.C.(A), G.C., Scint., Eml.
			1310 to 1721	7.5	
M-C	July 16	2320	0130 to 0132	18	I.C.(B), Eml.
			0924 to 0925	88	
IGC-16	July 17	1022	1223 to 1731	7	I.C.(A), Scint., Eml
			1830 to 2240	13	
M-D	July 17	2315	0137 to 0140	15	I.C.(B), Eml.
			0330 to 0810	25	
IGC-17	July 18	0531	0730 to 2000	8	I.C.(A), G.C., Scint., Eml.

I.C.(A) = ion chamber, IGY type.

I.C.(B) = ion chamber, continuous monitor chamber.

I.C.(C) = ion chamber, satellite unit chamber.

G.C. = Geiger counter.

Scint. = scintillation counter (NaI).

Eml. = nuclear emulsions.

increases in rates will be observed as the balloon proceeds upward even above the height of the normal transition maximum.

If the nature of the particles is known, then the variation of the intensity with depth may be used to evaluate the energy spectrum of the particles by converting a range spectrum to energy from known range-energy relations. Methods of carrying this out have been described and utilized in a recent paper describing the solar cosmic-ray event of May 1959 [Winckler and Bhavsar, 1960]. This procedure may be used only if the intensity above the atmosphere has a known angular distribution and is constant as the balloon rises, and if the nature of the radiation is known from some other means or by making reasonable assumptions. In any event, each flight yields the time and depth dependence of the total particle flux on ascent, and then at constant level shows the time variations occurring at constant depth. Some information about major changes in the shape of the spectrum at the balloon ceiling

height may be obtained by comparing the ionization chamber and Geiger counter in the manner described in connection with the May cosmic-ray event [Winckler and Bhavsar, 1960].

In most of the plots shown in this paper, the rates of the instruments have been averaged over 5-minute intervals during the constant level portion of the flight, but shorter intervals may be taken from time to time during ascent or during special rapid changes of intensity. A logarithmic scale is used for the Y axis because of the relatively large excursions in rates of the instruments during these events. For the ascending portion of each flight we have plotted separately the rates of the instruments against pressure on a log-log plot, which is useful for evaluating the energy spectrum of the particles. These plots also serve to show the depth at which the solar cosmic rays are first detected by comparing active days with normal days. The lowest energy to which the instruments are sensitive decreases, of course, as the balloon rises and the residual atmosphere decreases. At

TABLE 2. Characteristics of the Counting Instruments

Instrument	Wall Material and Thickness, mm	Geometry and Dimensions, mm	Omni- directional Projected Area, cm ²	Other Details
Ion chamber (A)	Steel 0.46	Sphere dia. = 254	506	Neher type pulsing chamber
Ion chamber (B)	Aluminum 1.30			1.90×10^{-10} coulombs/pulse.
Ion chamber (C)	Aluminum 0.51			Filling gas, Argon at 7.83 atmosphere.
Geiger counter	Brass end 1.32 side 0.79 end 4.20	Cylindrical Active length = 38.0 Active dia. = 49.2	24.17	Supplied by Wood Labs, self- quenching. Dead time $\approx 200 \mu$ sec
Scintillator (NaI (Tl))	Aluminum 1.47	Cylindrical height = 12.7 dia. = 45.45	12.64	Harshaw scintillation crystal.

Note: For the fast charged particles the number of ion pairs per second per cubic centimeter of standard air (76 cm hg 24°C) may be obtained from the normalized ion chamber rates, R_m (pulses/sec $\times 10^3$), according to the expression

$$N = 12.3 \times R_m$$

This is true for all the three types of ion chambers. For any other radiation the different wall thicknesses and other constants given in the table should be used.

the highest altitudes reached in this flight series, which correspond to between 6 and 10 g/cm², the minimum energy of the primary protons to which the instruments are sensitive is about 85 Mev. It is known that the solar-produced cosmic-radiation spectrum extends down to lower energies, perhaps as low as 30 Mev [Arnoldy, Hoffman, and Winckler, 1960b]. In the present paper, no direct information is available about the spectrum below approximately 85 Mev.

The scintillation counter, because of its high sensitivity of X rays and γ rays, shows large effects associated with auroras which accompany the magnetic storms and the solar protons at Minneapolis. Very high counting rates from X rays from the aurora were observed during this series of flights. The scintillator also responds to γ rays produced by nuclear interactions of the protons in the atmosphere above the balloon. Some discussion of the scintillation counter behavior was given in the paper describing the May cosmic-ray event [Winckler and Bhavsar, 1960]. Further discussion will be given by one of us (P. D. B.) about the interpretation of the scintillation counter measurements during auroras and solar proton events. For convenience a brief summary of the charac-

teristics of the instruments is given in Table 2. More details are given in the references indexed above.

The July 10 flare event. A class 3+ flare complex began as early as 0210 UT on July 10 and lasted for many hours with residual flare observations as late as 0900 UT. The first balloon flight in our series, IGC-12, reached high altitude at 0900 UT on July 10 (Fig. 2). No evidence of any disturbance was recorded on this flight, and normal cosmic-ray rates were observed throughout. The ionization chamber observations terminated owing to a malfunction at 1400 UT, but the Geiger counter rate was observed to be constant until 1930 UT. The normalized ionization chamber rate on this flight was approximately 18, characteristic of the galactic background during the period.

The next flight, IGC-13, reached ceiling altitude at 0518 UT on July 11 (Fig. 3). This flight showed some significant increases in ion chamber and counter during the ascent of the balloon, and further time variations during the level flight. The scintillation-counter-rate increase between 0800 and 1300 UT was caused by a balloon altitude decrease during this time. Considering first the ascent portion, Figure 4

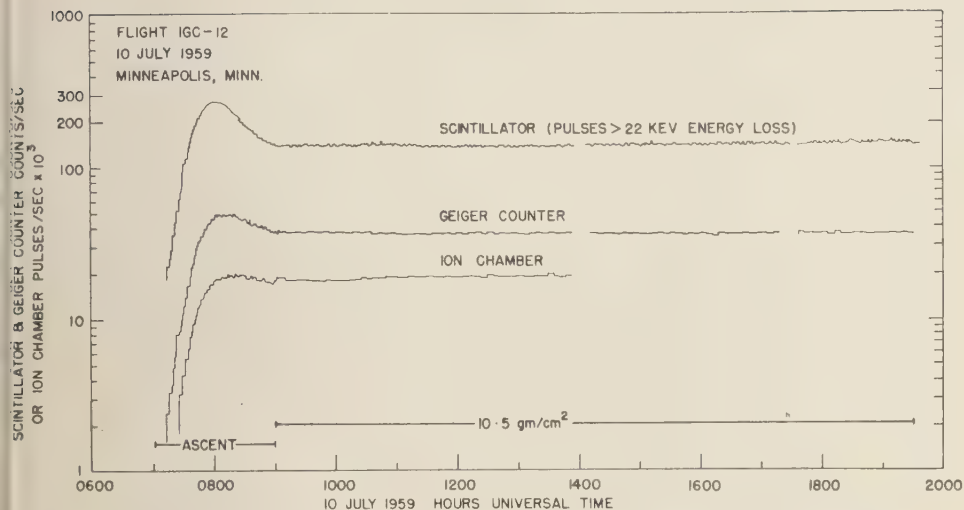


Fig. 2. Ion chamber, Geiger counter, and scintillator records between the time of the first flare at 0210 UT on July 10 and the geomagnetic storm, showing normal rates on all instruments characteristic of the galactic cosmic ray background.

shows the rates on both flights, IGC-12 and IGC-13, plotted against pressure for comparison. It will be seen that these flights agree very well at atmospheric depths greater than 100 g/cm^2 . Below this depth, flight IGC-13 departs, showing an increased radiation apparently having a transition maximum at about 35 g/cm^2 , and then the instrument rates decrease with decreasing depth until finally, at about 15 g/cm^2 , the ion chamber again shows an increase. Again,

referring to Figure 3, it is clear that this second increase is a time variation as the rates continue to climb after the balloon reaches ceiling. The peak normalized ion chamber rate is 36, and the counter rate, 46, compared with normal ceiling values of 18 for the ion chamber and 36.5 for the counter. This transitory increase slowly dies away during the flight, and at the time of parachute release at 1920 UT on July 11, the rates have essentially returned to

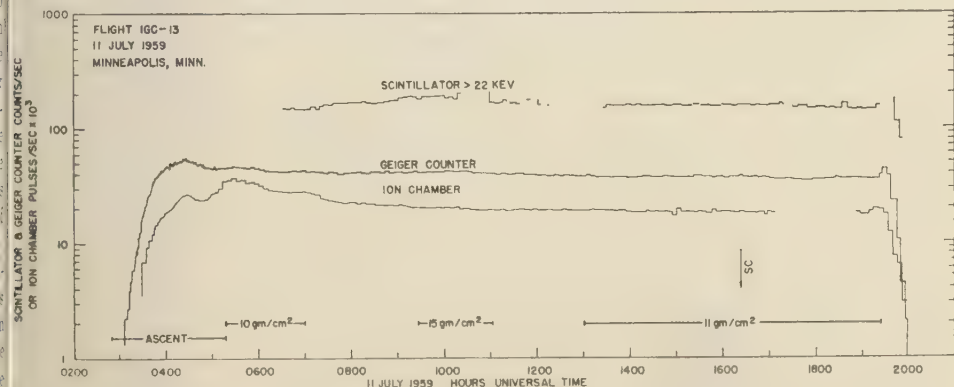


Fig. 3. Ion chamber, Geiger counter, and scintillator records showing the temporary influx of radiation between 0400 and 1000 UT on July 11 before the sudden commencement of the geomagnetic storm. Note also the lack of effect on any of the instruments at the time of the sudden commencement at 1623 UT.

normal (see Fig. 3). It is impossible to say with certainty whether the transition maximum observed on ascent at higher than normal rates is a time variation, a depth variation, or a combination of the two. At 0530 UT the peak excess radiation has a mean ionization 8 times that for minimum ionizing particles and an omnidirectional flux of 0.39 particles/cm²·sec.

The next series of events following the first flare is associated with the geomagnetic storm, and the flights are plotted in Figure 5. The sudden commencement occurred at 1623 UT on

July 11. Flight IGC-13 was at high altitude during this period, as can be seen from Figures 3 and 5. The change in the high-altitude intensity within 10 minutes of the sudden commencement is *less than 1 per cent*, and the normalized ion chamber rate is about 18.2 at the time of the sudden commencement, which is closely the same as its value the day before. At 1905 UT, 20 minutes before parachute release, IGC-13 showed a new rate increase. This increase was simultaneously shown by the next flight in the series, M-A (Fig. 5). The instrumentation on

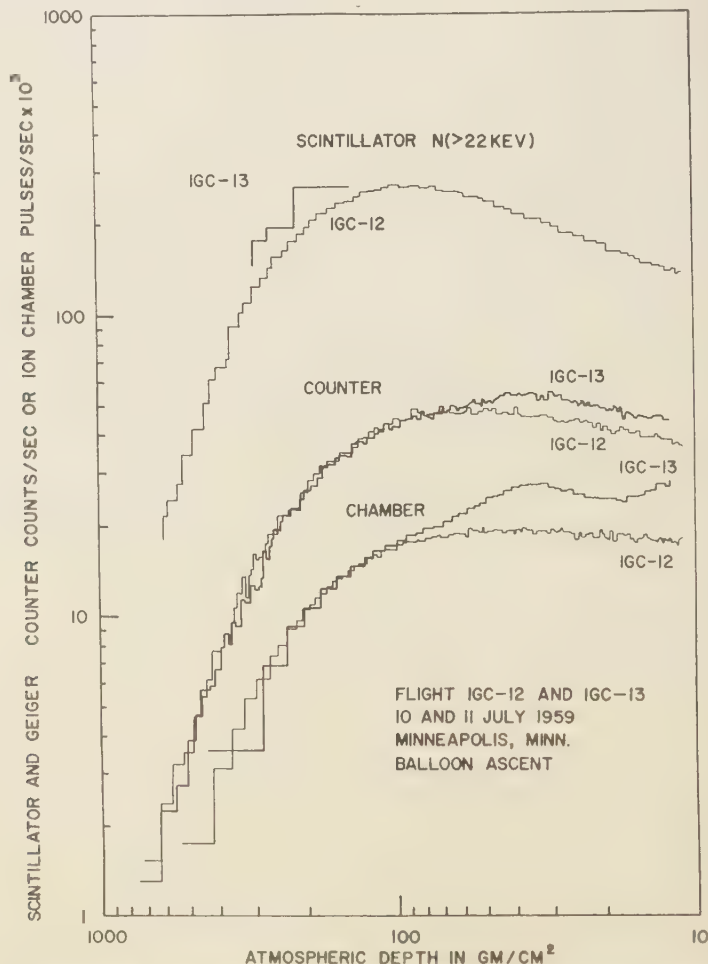


Fig. 4. The balloon instruments plotted against pressure during ascent of the undisturbed flight IGC-12, and the flight IGC-13, showing excess radiation before the geomagnetic storm. The excess radiation effect is a combination of a depth variation and a time variation, making the interpretation extremely uncertain.

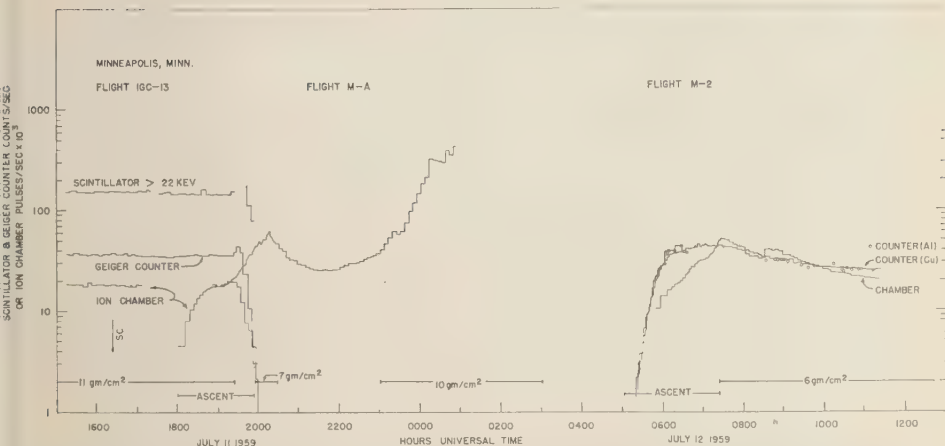


Fig. 5. Several balloon flights, showing the time dependence of the radiation during the geomagnetic storm period following the first flare. The first evidence of excess radiation during the storm occurred at 1905 UT, about 2.5 hours after the sudden commencement, but the main increase was delayed about 8 hours. The rates returned to near normal by 1100 UT on July 12.

M-A was a single aluminum ion chamber. This flight was launched at 1723 UT on July 11 and at the time the increase began had reached 27 cm^2/gm^2 atmospheric depth (Fig. 6). It continued to show an increase in rate until at 2015 UT it had reached 62 on the normalized scale. That this increase is a time variation can be easily seen by comparing Figure 5 with the rate-pressure plot in Figure 6. After this the rate began to decrease, reaching a minimum at 2145 UT, then began to climb until, at the end of the telemetry portion, it had increased to 430 on the normalized scale. In Figure 6 we show also the rate-pressure-versus-rate curve for the ion chamber on the ascent of IGC-12 for comparison. Although only the ion chamber was carried on flight M-A, the smooth character of the intensity changes labels these as probably associated with the solar cosmic rays and not with X ray bursts from the aurora. Flight M-A was still increasing in rate when the telemetry record ended. The value at this time was 24 times the ionization rate of the normal galactic background at the same altitude. Assuming protons with an energy spectrum proportional to $E^{-2.5}$ similar to that observed on previous events, the mean ionization for the particles would be about 7 times higher than the ionization due to the minimum ionizing particles. This means that the equivalent omnidirectional flux on July

12 at 0100 UT on flight M-A would have reached about 6.6 particles/ cm^2sec . Because of the steep spectrum, this equivalent omnidirectional flux is really concentrated in the vertical direction; the vertical particle flux would then have increased to approximately 66 times the 'normal' vertical particle flux of galactic cosmic rays

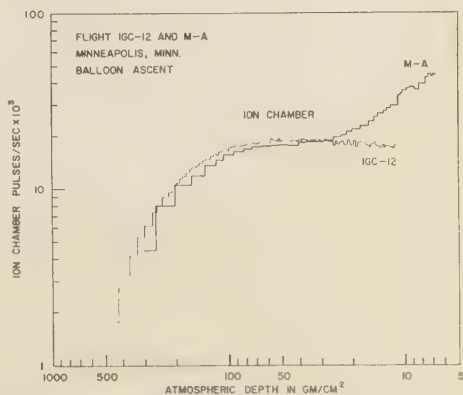


Fig. 6. Rate-pressure comparison on balloon ascent of an undisturbed flight, IGC-12, and a flight during the main phase of the magnetic storm, M-A. Note the decreased rates at great atmospheric depth associated with the Forbush decrease of the galactic cosmic radiation, and the increased rates at low depth associated with the solar cosmic rays. The intensity changes again are a combination of depth and time variations.

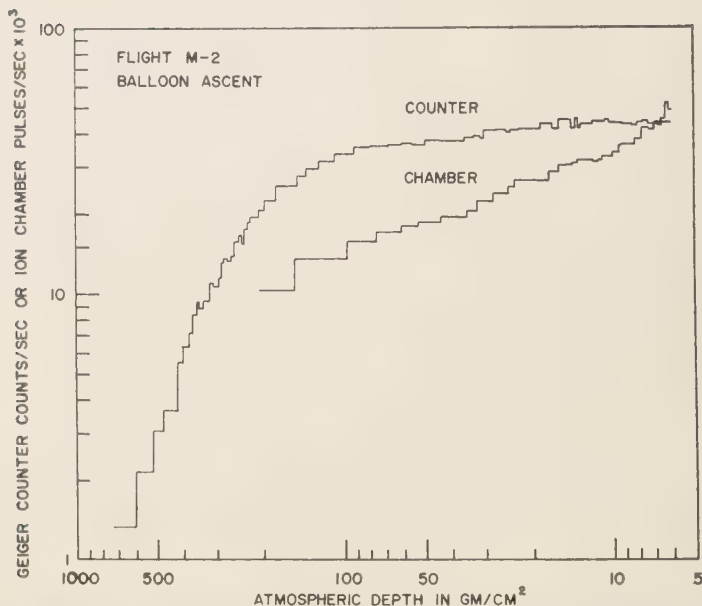


Fig. 7. Chamber and counter rate-pressure plot during the ascent of flight M-2, during the main geomagnetic disturbance of the first July event. In this case the increase of radiation with decreasing depth may be interpreted as showing that protons probably as low in energy as 80 Mev were present.

during the period of sunspot maximum. The rate was still on the increase when the telemetry ended, and so the peak value may have been considerably higher than the value at 0100 UT.

The final flight, M-2, reached high altitude at 0725 UT on July 12 (Fig. 5). At this time the rates were decaying from the previous increase shown on flight M-A. Flight M-2 was made with one of the first units of the new National Science Foundation continuous monitoring program, carrying aluminum and copper Geiger counters, ionization chamber, and vertical coincidence telescope. The last instrument did not operate properly, but we see in Figure 5 the records of both the copper and aluminum counters, which agree quite well, and the aluminum ionization chamber which has been normalized on the same scale as the previous type instruments. At the termination of the signal from this flight at 1120 UT on July 12, the ionization rate had dropped to 20, which is somewhat higher than galactic background, but most of the solar cosmic rays had disappeared at the time. The rate-pressure plot in Figure 7 shows

a continual increase of rate as pressure dropped to the highest altitude reached, approximately 6 g/cm^2 . This means that the spectrum of the particles, assuming protons incident at this time extended down at least to 85 Mev. We are assuming in this case that the increase with altitude is not a time variation, for when the balloon reached ceiling and floated level the time variation was shown to be of the opposite sense namely, a decrease of intensity with time. So although it is difficult to evaluate a spectrum from the rate-pressure curve of Figure 7, it is possible to set such a lower limit on the energy of the incident particles, assuming protons. Unfortunately, this procedure cannot be used in Figure 6 for it is fairly sure that the increase with altitude is compounded with an increase with time of the particle intensity. We see that during the event, at the time of the geomagnetic disturbance, the energy cutoff of the incident particles had dropped far below the value of $250 \pm 30 \text{ Mev}$ currently accepted for 'normal' times at Minneapolis [McDonald, 1957; McDonald and Webber, 1959; Fowler, Freier and Ney, 1958].

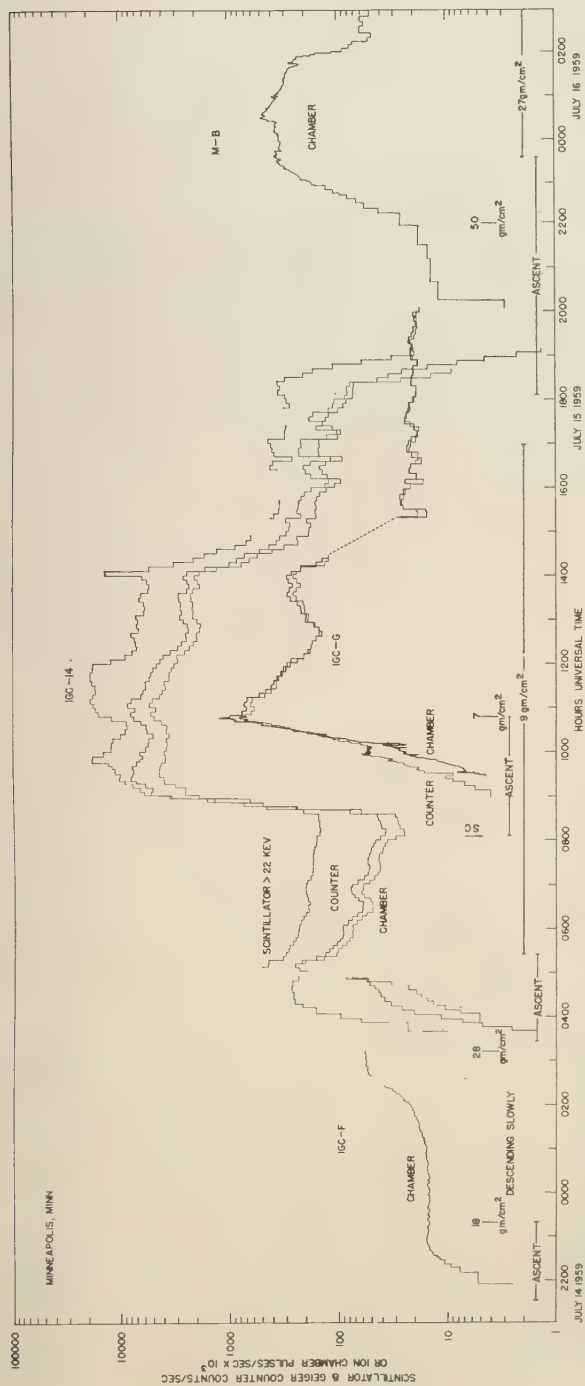


Fig. 8. Series of balloon observations following the second flare of the July series. A precursor effect occurred prior to the sudden commencement, shown by flights IGC-F and IGC-14, between 0000 UT and 0800 UT on July 14. Particles near geomagnetic cutoff apparently were present at this time. Note the rapid increase of radiation at 0830 UT, $\frac{1}{2}$ hour after the sudden commencement. This event is the most intense observed, in terms of ionization and counting rates. A simultaneous flight, IGC-G, is plotted on a scale one decade below that for IGC-14. Actually, the peak rates are in good agreement. (Time history continued on Fig. 11.)

The July 14 event. This event began with a class 3+ flare almost on central meridian as early as 0325 UT on July 14. This flare lasted at least 6 hours; one observation places the end at 1121 UT. Again this seems to be a flare complex with possibly two separate maxima at 0349 UT and 0527 UT, reported by separate observatories [*National Bureau of Standards*, 1959]. The first balloon flight after the flare, IGC-F, was launched at 2130 UT on July 14. This balloon reached a ceiling altitude of 18 g/cm² at 2320 UT (Fig. 8).

On reaching ceiling, the ion chamber—the

only instrument on this flight—showed a value of 15 compared with the value of 18 on July 11, before the commencement of the first magnetic storm during these events. This low rate of 15 remained constant up to about 0100 UT on July 15, when an increase began. It reached a value of about 60 at 0312 UT, at which time the signal was lost. Because of some balloon fault, this flight did not float at a constant altitude after reaching the ceiling, but descended steadily, and was at 28 g/cm² atmospheric depth when the signal was lost. It is interesting to note that the rates were increasing even though

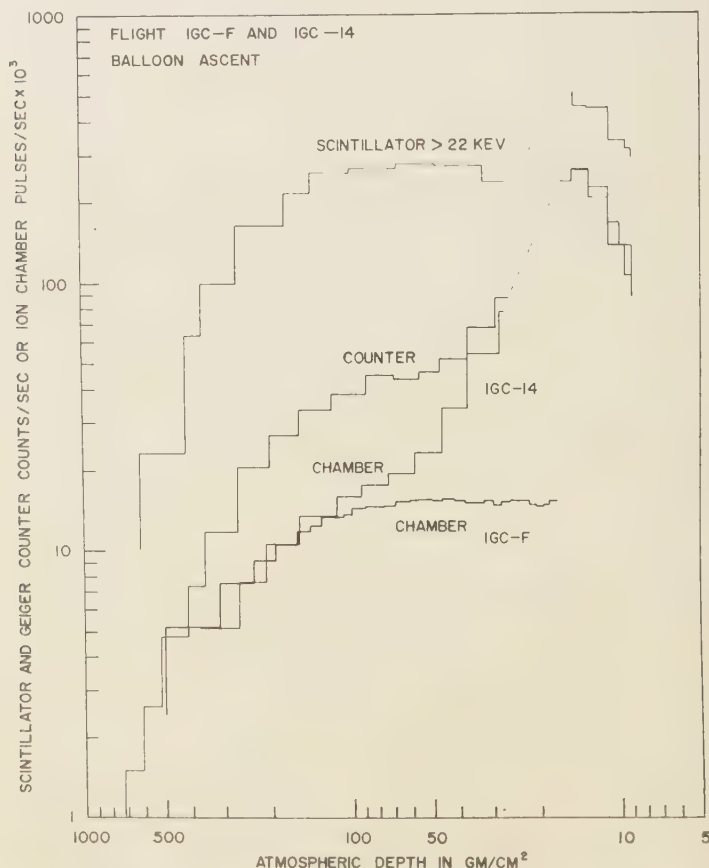


Fig. 9. Rate-pressure plot for the radiation observed before the geomagnetic storm for the second July event. The presence of the radiation is clearly shown by comparing the ion chambers on IGC-F and IGC-14. Some energy spectral information can be obtained from the data in this figure, and protons down to about 200 Mev seem to be present. However, the scintillator rates do not show a great intensity of γ radiation from the intense low-energy portions of the spectrum above the balloon altitudes.

the balloon was descending. This behavior must therefore be a time-variation effect. It also suggests that the particles arriving at this time must have had a range larger than 28 g/cm^2 i.e., of $KE > 200 \text{ Mev}$ for protons). They were arriving well in advance of the sudden commencement which occurred a few hours later that day.

The next flight in the series, flight IGC-14, was launched at 0325 UT on July 15 and carried a complete complement of instruments, including the scintillation counter. This flight was launched from the usual site near Minneapolis at sunset and the giant spot group in which the flare had occurred could be seen easily with the naked eye on the red disk of the sun near the horizon. This flight first detected excess radiation at approximately 100 g/cm^2 , and the radiation intensity increased in the ionization chamber to a value of about 250 on the normalized scale when the balloon reached an altitude of 4 g/cm^2 . After this the intensity decreased to a value of 107 as the balloon reached the ceiling altitude of 9 g/cm^2 at 0525 UT. In the next hours the radiation intensity decreased until on the normalized scale it was about 30, which is only about twice normal galactic intensity. We see thus that the flights covered what appears to be another precursor effect ahead of the geomagnetic storm, in this case with considerable intensity. The sudden commencement of the magnetic storm occurred at 0802 UT on July 15 as indicated in Figure 8. There appears to be a small initial drop in the intensity of the solar cosmic radiation at the time of the sudden commencement, but this effect is only about 20 per cent of the excess intensity above galactic radiation and is not greater than other fluctuations seen earlier on the flight; however, it may perhaps be significant. Intensities then increased somewhat until 0835 UT. Then between 0835 UT and 0900 UT, approximately, a very spectacular and rapid rise in rate occurred with a total factor of over 100 and with the normalized ionization rate reaching over 5000. This increase occurred with the balloon floating at ceiling and is certainly a sudden and dramatic time variation following the sudden commencement of the magnetic storm by $\frac{1}{2}$ hour. The ionization counter and scintillator follow very closely the same relative pattern. This high intensity persisted for about 5 hours with fluctuations

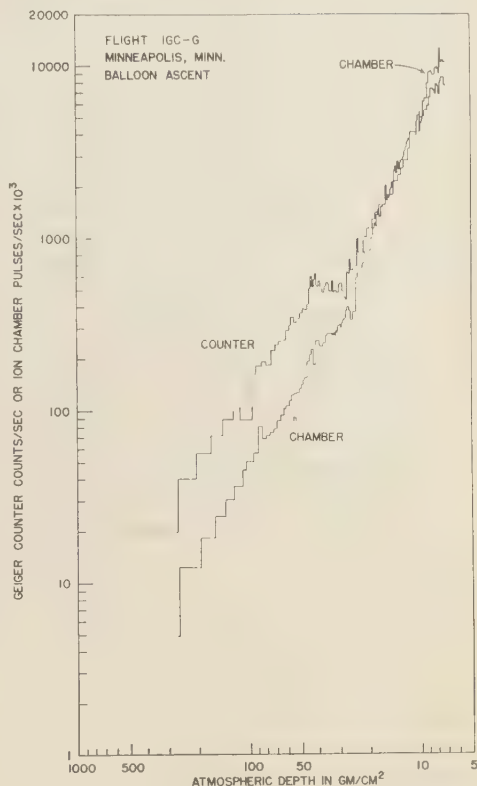


Fig. 10. Rate-pressure plot during ascent of IGC-G. The ion chamber and Geiger counter were designed especially for high rate conditions for use in the satellite and space probe series Able-5 of the NASA. This flight ascended while flight IGC-14 monitored the intensity at high altitude. It is the best flight of the series for unfolding the range spectrum of the solar protons in the atmosphere (see text). The dosage rate at 7 g/cm^2 is approximately 0.15 r/hr .

ranging from 5000 to 2000 on the normalized scale. The intensity then began to drop rather fast and reached a plateau value of 150 to 250 on the normalized scale at the time of balloon parachute release at approximately 1820 UT. The rates dropped as the parachute descended to the ground. This high-intensity period was also a period of auroral activity, and the scintillation counter showed definite increases not shown on the other instruments as evidence for auroral X rays.

A further flight, IGC-G, was launched at 0804 UT on July 15 and ascended during a 5-hour period of high fairly uniform intensity.

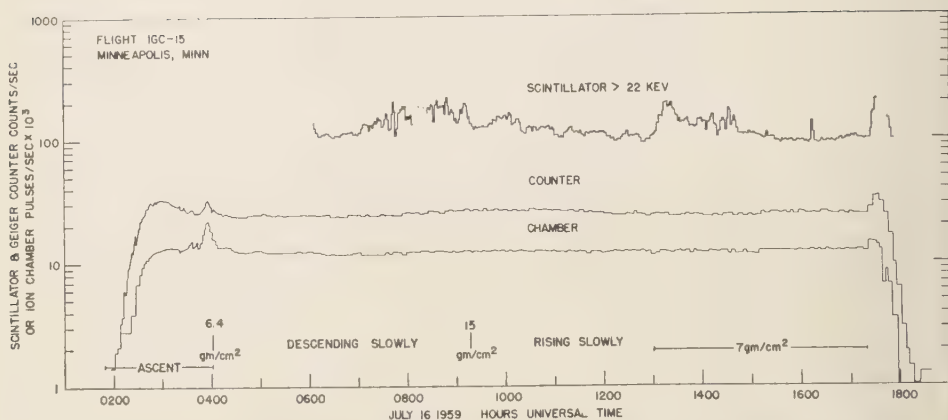


Fig. 11. Ion chamber, counter, and scintillator records at the end of the second July event. The increase at 0400 UT is probably an X-ray burst associated with auroral activity. The scintillation counter record available after 0600 UT shows continued evidence for weak auroral X-ray activity. There is little or no evidence for solar cosmic rays on this flight, and the event therefore ended by 0300 UT on July 16 at Minneapolis.

The instrumentation consisted of an aluminum ion chamber and counter of the types designed for the Able satellites and space probes for the NASA, which were flown on Explorer VI earth satellite and Pioneer V space probes [Arnoldy, Hoffman and Winckler, 1960a]. These instruments are in principle the same as the standard ionization chambers and counters but have different sizes and sensitivities. Figure 8, also shows the record from flight IGC-G for the ascending portion. In this figure the rate scale for IGC-G has been displaced downward one decade to avoid confusing the curves for the two flights. Flight IGC-G reached a somewhat smaller atmospheric depth, but in general the rates are in excellent agreement on the normalized basis with IGC-14. The two balloons continued to drift in a westerly direction from Minneapolis, with IGC-14 several hundred miles west of IGC-G. The rapid decrease between 1400 and 1500 UT is also shown on both flights but with some data missing from IGC-G during that interval. Flight IGC-G remained aloft longer and showed that the plateau value after the high-intensity period persisted at about 200 on the normalized scale until the end of IGC-G at 2000 UT on July 15. During the period 1500 to 2000 UT, very interesting semiperiodic fluctuations occurred on both flights which will be discussed below.

The next flight, M-B, was launched at 1807

UT on July 15. The time-rate plot for this flight is also shown in Figure 8. A rapid increase in radiation occurred between 50 g/cm² and 27 g/cm² during balloon ascent. The increase has the appearance of a solar proton increase, but since this flight had only an aluminum ion chamber, auroral X rays may be a contributing factor.

The final flight in the series covering the second July event, IGC-15, was launched at 0148 UT on July 16. Both the time-rate plots for this flight (Fig. 11) and the pressure-rate plot (Fig. 12) show that most of the ascent was normal with the exception of a burst of radiation as the balloon reached ceiling altitude which appears to be X rays from auroral activity. Unfortunately, the scintillation counter record is not available for this portion of the flight but begins at 0600 UT (Fig. 12). This instrument shows throughout the flight the irregular increases characteristic of fairly weak auroral X rays that are not detectable, however, on the ionization chamber and counter, probably because their energy and intensity were too low. It can be stated, therefore, that the solar cosmic-ray event at Minneapolis was ended, at the latest, by 0330 UT on July 16. By this time the magnetic indices as shown in Figure 1 had decreased to a fairly low value. It is interesting that the galactic cosmic-ray intensities measured on flight IGC-15 had dropped to about 1

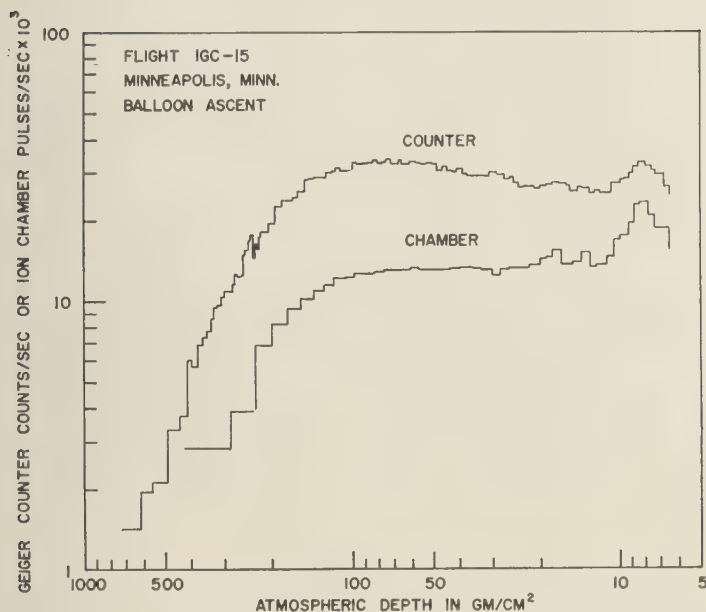


Fig. 12. Rate-pressure curves for flight IGC-15 that show little or no evidence for solar protons but do show an increase that may be auroral X rays appearing at high altitude.

on the normalized scale, compared with approximately 18 before the July series of events began. This 30 per cent decrease in the high-altitude ionization at approximately 10 g/cm^2 is to be compared with the approximately 16 per cent decrease in the sea-level neutron monitor as plotted on Figure 1.

This second event of the July series is probably the most interesting of all in relation to the purpose of this paper—to discuss the time variations and their correlations. Again in this event we see a precursor of intensity increase occurring before the beginning of the magnetic storm but 20 hours after the great flare on July 14. As in the first event of the period, no very large effects occurred at the time of the sudden commencement of the magnetic storm, but the great increase occurred *after* the sudden commencement, in this case with a time lapse of about $\frac{1}{2}$ hour. The sudden admission of the full solar cosmic-ray intensity at the geomagnetic latitude of Minneapolis resulted in a spectacularly high flux, the highest that has ever been observed at balloon altitudes in the course of all observations made by the Minneapolis group. The normalized ion-cham-

ber rate reached a peak value of 5800 on the normalized scale at 1100 UT on July 15. At these rates a partial saturation of the telemetry occurred, so that the observed peak ionization rates during IGC-14 may be 50 per cent lower than the actual rates. From the flight IGC-G rate-pressure curves, however (shown in Fig. 10), and assuming an isotropic incident flux of protons, the energy spectrum may be evaluated with some confidence as saturation did not occur. It can be represented by the following equation

$$N(>E) = 1.05 \times 10^8 E^{-2.9} / \text{cm}^2 \cdot \text{sec} \cdot \text{ster}$$

This equation applies above the atmosphere with E expressed in Mev above 88 Mev but is not accurate for $E > 300$ Mev. The total flux above the atmosphere greater than 88 Mev is then

$$N(>88 \text{ Mev}) = 241 / \text{cm}^2 \cdot \text{sec} \cdot \text{ster}$$

at 1046 UT on July 15. The continued increase of intensity as flight IGC-G ascended to a minimum depth of 7 g/cm^2 , equivalent to 88-Mev proton range, shows the large decrease in geo-

magnetic cutoff energy from the normal value of 250 Mev.

This event also offered an opportunity to determine the spectrum of the protons arriving before the sudden commencement from the rate-pressure plots of IGC-14 and IGC-F shown in

Figure 9 and the time-rate plot of IGC-F shown in Figure 8. When the signal was lost on IGC-F at 0315 UT, the ion chamber rate was 60 and the balloon altitude was 28 g/cm^2 . The rate-pressure curve for IGC-14 shows about the same rate for its ion chamber at the same

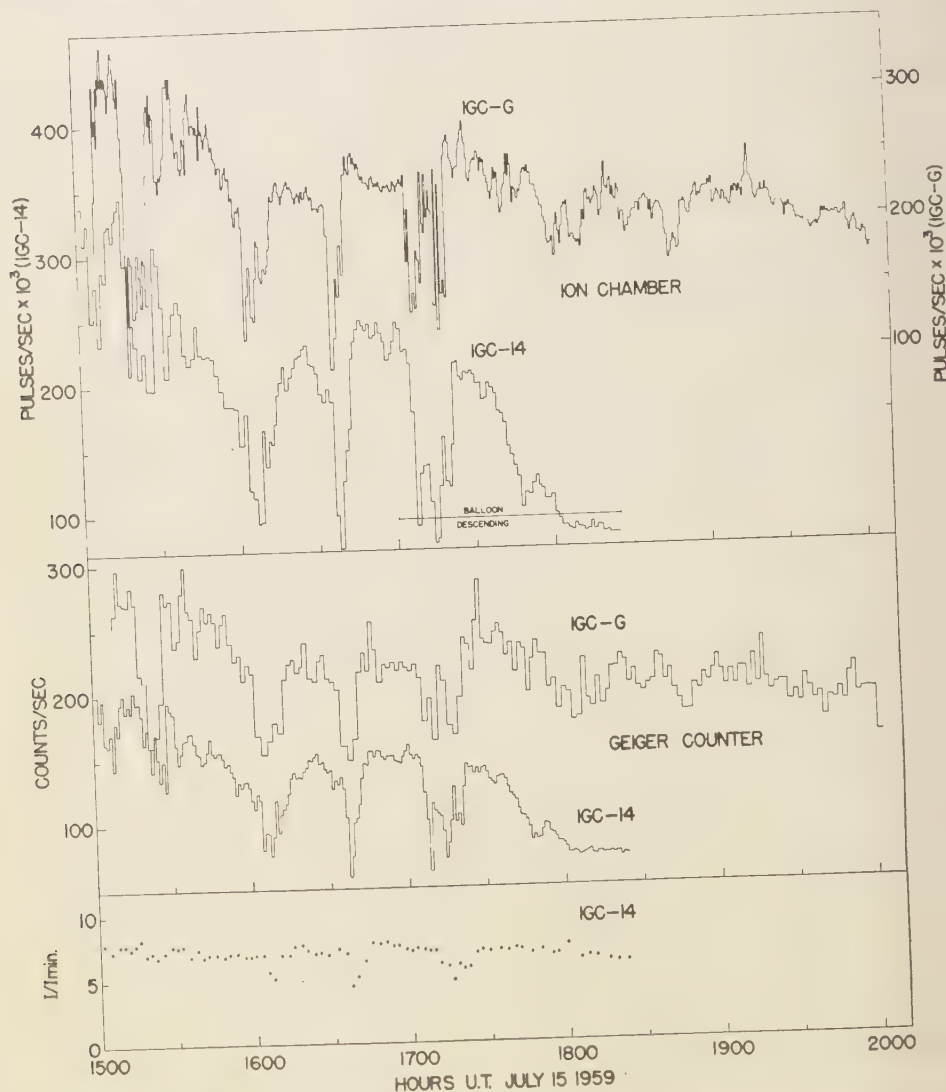


Fig. 13. Unusual semiperiodic oscillations in the solar cosmic-ray intensity observed near the end of the second July event. The records were made simultaneously with ion chambers and Geiger counters on two separate balloons about 200 miles apart near Minneapolis. There is some change in the character of the spectrum, as shown by the I/I_{\min} ratios (lower part of figure) during the sharp intensity drops.

atmospheric depth of 28 g/cm² 2 hours after IGC-F observation. If we assume from the above evidence that the excess proton intensity above the atmosphere was constant during this period, then we can determine the exponent of the proton spectrum from the rate-pressure plot on flight IGC-14 during ascent. To determine the excess number of protons on IGC-14 the C-F ascent can be used for quiet-day rate corrections, as shown in Figure 9. The spectrum determined in this manner for the protons arriving before the sudden commencement can be represented as

$$N(>E) = 5.1 \times 10^{12} E^{-5.5} / \text{cm}^2 \cdot \text{sec} \cdot \text{ster}$$

for protons, with $KE > 200$ Mev, at 0500 UT on July 15. This gives the proton intensity at that time as 1.0 protons·sec⁻¹·cm⁻²·ster⁻¹ of energy above 200 Mev at the top of the atmosphere.

The above-mentioned direct derivations are from the observations extending up to 28 g/cm² atmospheric depth, which limits the energy to about 200 Mev minimum. The scintillation counter, which is sensitive to the γ rays produced at smaller depths by protons of energy below 200 Mev, offers, however, a means to estimate their intensity. On flight IGC-8 in May 1959 [Winckler and Bhavsar, 1960], during a similar event above the proton spectrum and intensity were about the same as at this time, the scintillation counter was recording about 10 times more γ rays than in the present event. This may be taken as indirect evidence that the low-energy part of the spectrum was depleted, possibly by cutoff before the sudden commencement.

At the end of this paper, we discuss the possible intensity of the solar protons in free space near the earth deduced from these balloon flights and from recent measurements taken on April 1, 1960, simultaneously with the NASA deep-space probe Pioneer V and balloon flights. An examination of the various figures showing the ionization, count rate, and scintillation rate for these events shows a multitude of minor variations in addition to the major trend of intensities. Figure 13 gives in detail what is perhaps the most interesting of all these smaller variations in intensity that occurred near the end of the July 15 event as seen at Minneapolis. This record, between 1500 and 1900 UT on July 15, compares two balloons floating at high altitude approximately 200 miles apart in an E-W

direction near Minneapolis. The rates on the average remain at a fairly constant level but show many fluctuations that appear to be almost periodic and can be seen in Figure 13 with maxima at 1515, 1545, 1625, 1655, and 1730 UT. The maxima are quite flat with sharp drops in intensity between that show on all of the instruments on both balloons. Figure 13 also plots the mean ionization/count expressed in units of minimum ionizing particles. The ratio drops with decreasing intensity, showing that the changes are partly spectral in nature and not a pure intensity phenomenon.

The periodic nature of this variation presents an interesting and baffling problem. Plasma oscillations in the dipole field of the earth have periods much shorter than the approximately 30-minute period of these oscillations. The 30-minute period is approximately the precessional time of 30-Mev trapped particles in the earth's field. Although we are not sensitive to particles of this energy directly, the solar cosmic-ray spectrum does include such energies. The energy density would be too low, however, to produce an appreciable modulation of the earth's magnetic field. A detailed examination of the magnetograms during the period of these fluctuations does not show any corresponding variation. The phenomenon may very well arise in the solar beam in space. If suitable field configurations exist in space—for example, an extension of the coronal field carried out during the storm to the earth—then a 30-minute period might be due to 100-Mev solar protons moving in groups back and forth along this field between reflecting regions. Kellogg (private communication) has suggested that oscillations of this period may follow a hydromagnetic shock front arriving from the sun, and that this may modulate the injection of the solar cosmic rays into the geomagnetic field.

The July 16 event. The last of the great events began with a 3+ flare on July 16 at 2114 UT in the same complex spot group as the other two flares. Flight M-C was launched at 2320 UT on July 16 and reached high altitude at 0130 UT on July 17 (Fig. 14). This flight was instrumented with a single ionization chamber and showed the presence of excess radiation at atmospheric depths below about 50 g/cm². There is also evidence that the increase of radiation with decreasing depth was coupled with a

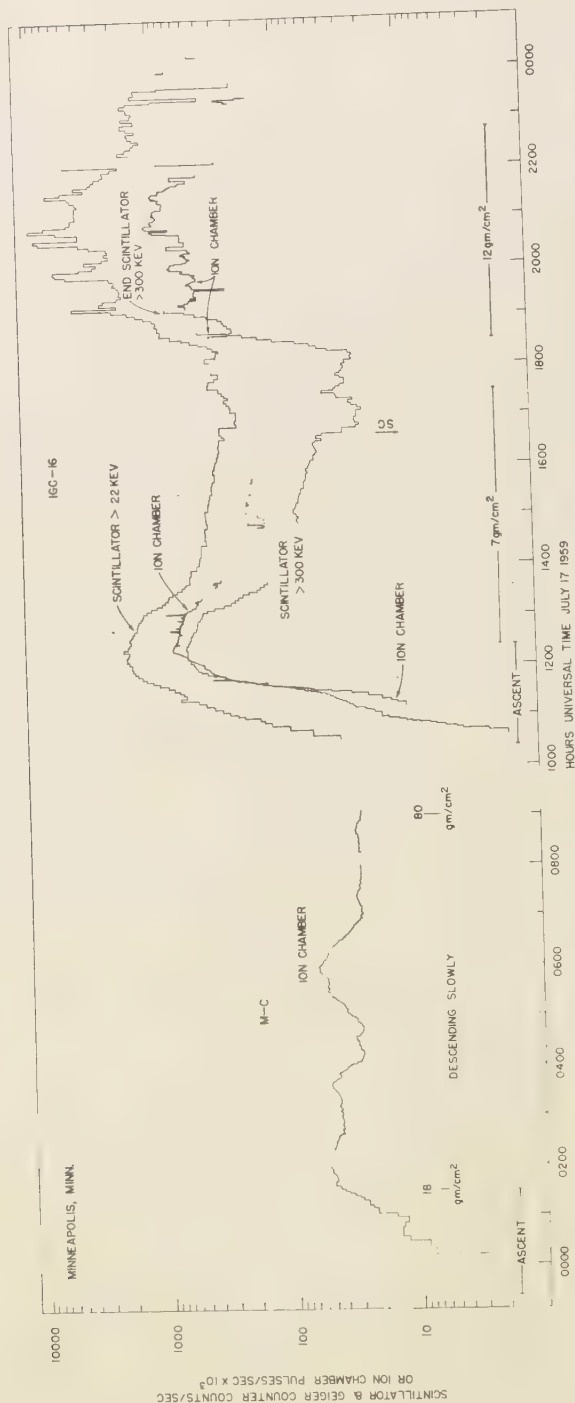


Fig. 14. Recordings of balloon instruments before and during the geomagnetic storm associated with the third July event. Note the sudden commencement at 1640 UT on July 17. On this date, before the geomagnetic storm, the solar cosmic rays were observed at sea level, as well as the effects shown above at high altitude. The scintillator channel >300 kev responds mainly to charged particles and represents approximately the Geiger counter response. In addition to proton increases, auroral X-ray activity occurred during the main geomagnetic storm after 1800 UT, as shown by the spikes on the scintillator record at that time. Note the substantial drop in intensity at the time of the sudden commencement.

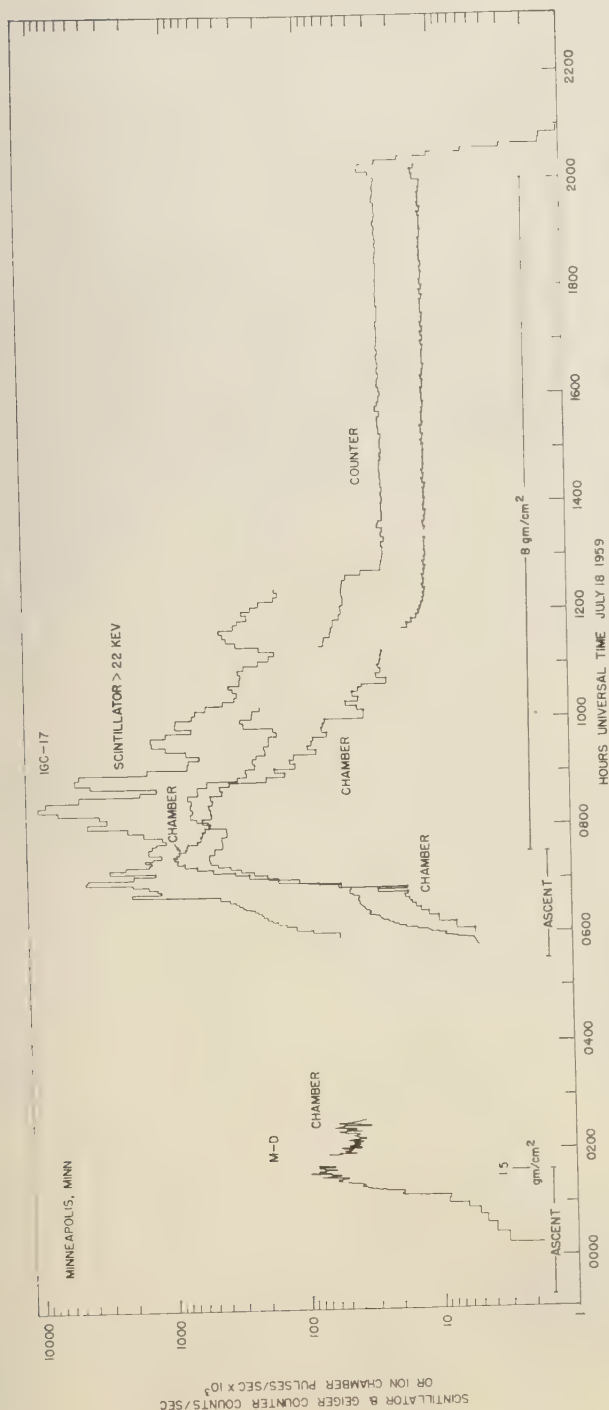


Fig. 15. Final part of the Minneapolis cosmic ray increase for the third July event. The auroral X-ray activity was superposed on the solar proton increase, under way when the balloon reached ceiling at 0730 UT. Only galactic cosmic-ray rates were obtained by about 1300 UT, but at this time a large flux of protons was still observed in the polar region.

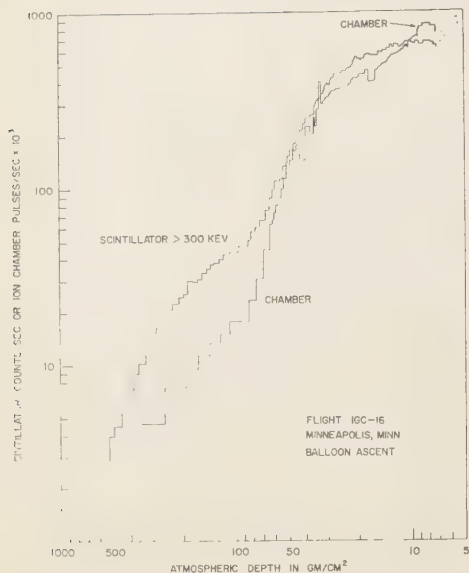


Fig. 16. Rate-pressure curves for the ascent of flight IGC-16. The curving over of the intensities is a combination of the decreasing depth of the balloon and the decreasing intensity of the radiation above the atmosphere.

time variation as the rates continued to climb after the balloon reached peak altitude. The normalized ionization rate reached 56 at 0210 UT on July 17 and remained at more or less this same high level until the end of the flight at 0900 UT. The considerable fluctuations that occurred are shown in Figure 14.

Flight IGC-16 was launched at 1022 UT on July 17 and showed a still higher level of radiation detectable at an atmospheric depth of about 100 g/cm². The rate-pressure curves for IGC-16 and IGC-17 are shown in Figures 16 and 17. By comparing Figures 14 and 16 it can be seen that the balloon flight IGC-16 was ascending at a time when the radiation intensity was dropping, and that when the balloon reached ceiling the rate was 850 on the normalized scale and immediately began to fall. Flight IGC-16 included, besides the ionization chamber, a scintillator with two channels, one looking at pulses > 300 kev, which means that this channel is reading particle flux essentially similar to the Geiger counter. The other channel on the scintillator read all pulses > 22 kev, which makes its response sensitive to photons

of the type found in the auroral X rays. Although the ionization chamber record is incomplete for flight IGC-16, it can be seen by referring to the high-energy scintillator channel that a substantial drop in the solar cosmic-ray intensity occurred at the time of the sudden commencement at 1638 UT on July 17 (Fig. 14). At about 1815 UT a large increase occurred that was similar in many respects to that for the July 15 event (flight IGC-14). The peak ionization rate reached 1200 on the normalized scale at 2050 UT. The record on this portion fluctuates a great deal, which is probable evidence for strong bursts of auroral X rays superposed on the solar cosmic rays. The intensity then began to fall, and the next flight in the series shown in Figure 15, flight M-D, shows a decrease in intensity which had a maximum value of 100 on the ionization scale at 0130 UT on July 18. A further strong rise in intensity was observed at 0700 UT on July 18 on flight IGC-17 (Fig. 15). The ionization rate again climbed

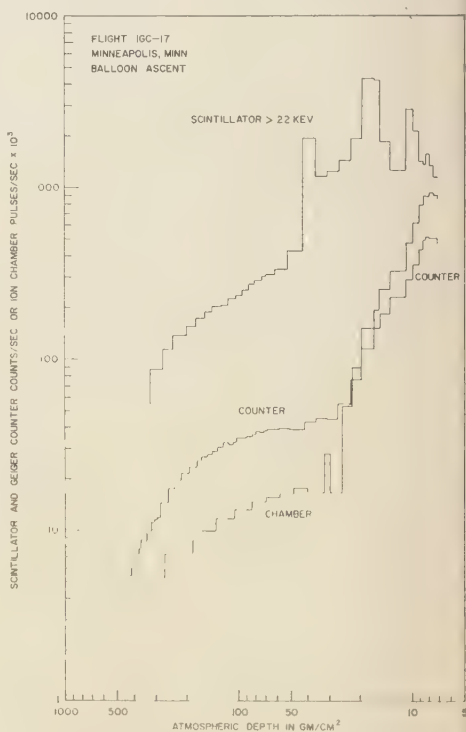


Fig. 17. Auroral X rays and solar cosmic rays shown on the ascent of flight IGC-17.

about 900 on the normalized scale. The rate then fell steadily and rapidly until at 1230 UT on July 18 it had reached galactic cosmic-ray background with no evidence at all thereafter for solar cosmic radiation up to 2000 UT. The galactic intensity in terms of the normalized high-altitude ionization rate at this time reached about 11.2 on the normalized scale, the lowest value that has been observed at Minnesota during the period of sunspot maximum. The scintillation counter on flight IGC-17 showed evidence for much auroral activity and large increases due to auroral X rays, as on many of the other flights.

The third event in the July series is interesting because of the rapid appearance at low geomagnetic latitude of an effect from the solar cosmic rays following the flare, and because during the period before the sudden commencement there was evidence for an effect at sea level. The sea-level effect as analyzed by Carmichael [Carmichael and Steljes, 1959] is reproduced in Figure 18. This analysis shows that high-altitude and particularly high-latitude neutron monitors showed a definite increase that was not seen at low latitude. The interpretation is clearly

that of a small but definite sea-level effect due to the solar-produced cosmic radiation from this event. This implies that the rigidity spectrum extended to about 1 bv even though the low-energy content was not greater than that observed in other events of the July series.

The rapid decrease of the cosmic-ray intensity after the last flare of the July series follows the decay of magnetic activity shown by the K -indices in the magnetic storm period (Figs. 1a and 15). In contrast, measurements made with balloons show that the high-latitude solar cosmic-ray flux was maintained for many days following this event. These measurements were made at Churchill by J. Earl (University of Minnesota, unpublished; see Fig. 1a), and at Resolute Bay by Anderson and Enemark [1960]. Anderson's measurements detected the solar cosmic rays for 10 days after the third flare.

Detailed analysis of the storm field and the changes in cosmic ray rigidities. The evidence indicates that in most cases the large changes of ionizing radiation intensity at Minneapolis during the storms are not a result of real intensity fluctuations in the beam in external space but are associated only with the environment of the

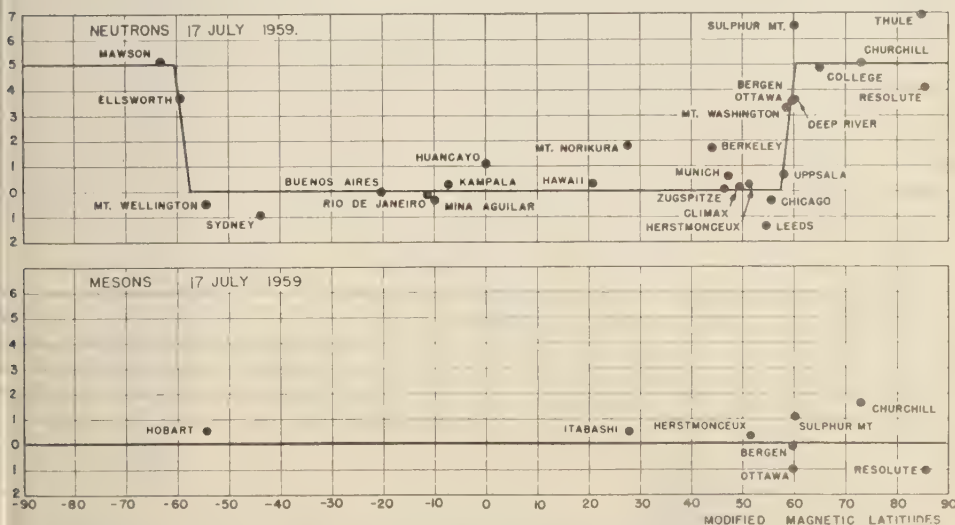


Fig. 18. Analysis of sea-level cosmic-ray neutron monitors by Carmichael for the third event on July 17. The energy spectrum extended to values such that high-latitude sea-level monitors in both hemispheres were activated. Little or no effect was noticed on the meson component.

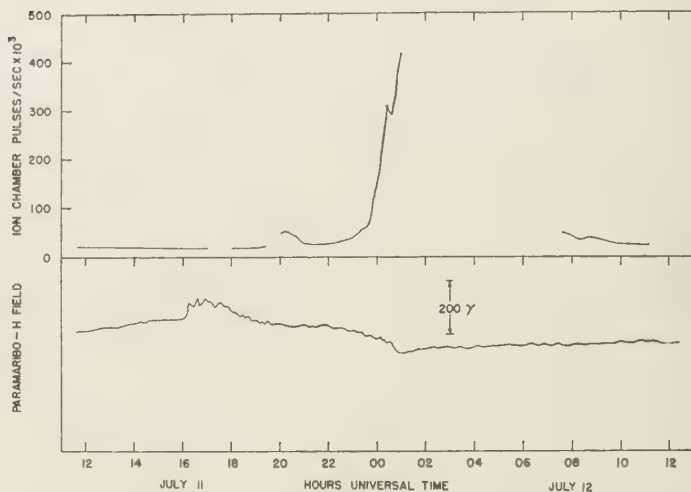


Fig. 19. High-altitude ionization rate, above, and equatorial horizontal field of the earth, below. The major increase at Minneapolis occurred when the equatorial field changed from positive to negative at the beginning of the main phase of the storm. Note the approximately 8-hour interval between the sudden commencement and the main phase.

earth. We should now like to make a detailed examination of the intensity increases in correlation with the changes in the main field of the earth. In Figures 19, 20, and 21 we have plotted the equatorial field of the earth measured at Paramaribo during the three storms and, for correlation, the ionization rate measured at

Minneapolis during the constant level portion of all available balloon flights. These three figures show that, although there may be precursor effects preceding the main event, the principal increase of cosmic-ray intensity occurs when the horizontal field of the earth passes negative into the main phase of the magnetic storm. We do

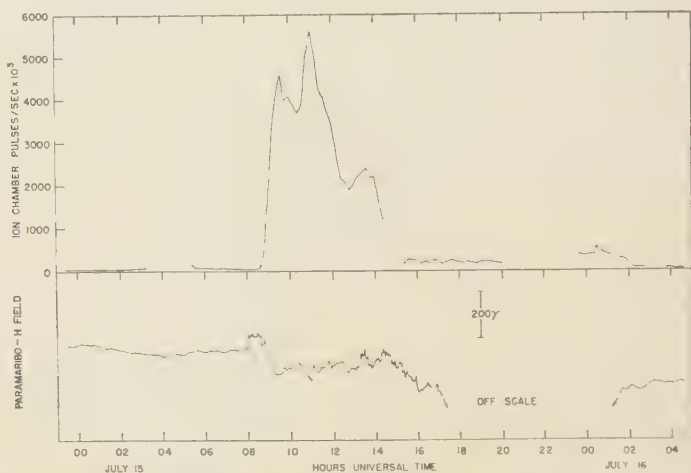


Fig. 20. High-altitude ionization rate, above, and equatorial horizontal field of the earth, below, during the second geomagnetic storm of the July series. Note again that the major increase at Minneapolis occurred when the storm field entered the main phase, in this case only about $\frac{1}{2}$ hour after the sudden commencement.

to see effects at the time of storm commencement, e.g., in Figure 19 the sudden commencement occurs at 1612 UT with no measurable cosmic-ray effect. The main phase of the storm occurs at about 2300 UT, and at this time a large increase in the intensity of the solar cosmic radiation begins. There are no balloon flight data available from 0100 until about 0800 UT on July 12, but then the intensities have returned to normal. In Figure 20 we see again that the sudden commencement occurs at 0804 UT on July 15 but no detectable effect occurs in the cosmic-ray intensity. When the main field of the earth goes negative about 30 minutes later, at about 0835 UT, the great increase of cosmic radiation begins. In this case the intensity is well covered by flights and remains high until at least 0000 on July 15. The third event on July 17 (Fig. 21) shows a major increase at 1200 UT before the main phase. As discussed in the previous section, this is believed to be a high-energy event in which the cosmic-ray particles are actually incident with energies above the normal cutoff at Minneapolis (see also Fig. 18). The sudden commencement occurs at about 1620 UT on July 17, but the major increase of cosmic rays is shown, by a balloon flight, to start at about 1900 UT, which coincides quite well with the beginning of the main phase of the third event.

The last balloon flight of the series, which also shows a large time variation, does not appear to

be correlated with major changes in the magnetic field at the Paramaribo station, but the storm was still in the negative phase. No balloon data are available following the next positive excursion and sudden commencement at about 1530 UT on July 18 (not shown in Fig. 21). There is thus a strong suggestion that the initial change of the field from the positive to the negative phase is associated with the lowering of the geomagnetic cutoffs, even though the positive phase may vary in duration from 30 minutes (Fig. 20) to 8 hours (Fig. 19). Other events besides those discussed in this paper are consistent with this same interpretation. In particular, events on March 26, 1958 [Winckler, 1960], and on May 12, 1959 [Winckler and Bhavsar, 1960], can likewise be interpreted as the lowering of cutoffs at the time of the main phase. Other events were observed in April, May, June, and September of 1960 that show similar correlations with the local magnetic field at Minneapolis. These events are being studied and the equatorial magnetograms will be examined when they become available.

It must be clearly stated that although direct spectrum measurements from many sources during the main phase show protons far below normal cutoff energies at Minneapolis, the *total absence* of such particles before the main phase *cannot* be established using the results of the counting instruments flown on our balloon flights. We have computed that for 10 g/cm^2

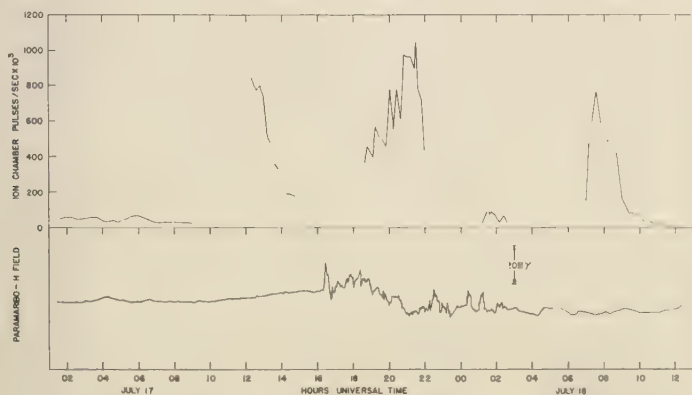


Fig. 21. High-altitude ionization rate, above, and equatorial horizontal field of the earth, below. The high altitude increase at about 1206 UT is probably high-energy particles above normal cutoff energies (see Figs. 1a and 18). The major increase occurs again when the storm field changes to the negative phase. A further increase occurred during the negative phase beginning at 0700 UT on July 18.

depth, and for 'normal' cosmic-ray cutoff conditions of approximately 250 Mev [McDonald, 1957; Fowler, Freier and Ney, 1958; McDonald and Webber, 1959], the maximum ratio of ions/count minimum for the solar proton component would be about 2. Any value appreciably higher than twice minimum suggests the presence of protons below normal cutoff at Minneapolis. In fact, some flights show moderate increases before the main phase, or even before the sudden commencement, which show ratios minimum in the range of 4 to 8. See, for example, Figure 3, 0530 UT July 11, with ratio 8; Figure 8, 0400-0800 UT July 15, with ratio 4; Figure 14, 1230 UT July 17, with ratio 5.3.

Two assumptions can be made about the prestorm radiation: (a) it includes α particles or heavy nuclei which increase the mean ionization; (b) Liouville's theorem does not apply, so that despite the lowered cutoff for some particles, free space intensities are not obtained on the earth. This condition could be due to time-varying magnetic fields. But it seems clear that a flux of particles as great as that in free space is observed at our latitude only at the start of the main phase.

The data also show that the cosmic-ray intensities return to normal before the end of the main phase of the magnetic storm. This is evident in Figures 19, 20, and 21, but it is more clearly shown by Figure 22. Here we have plotted the *Dst* curve for *H* for the three storms as determined by Akasofu and Chapman (private communication) from averages of a number of equatorial magnetic observatories corrected for the normal diurnal variation during quiet periods. We have plotted on the same time scale, in the upper part of the figure, the balloon ion chamber rates at high altitude. Because of the compressed time scale this curve is not so useful for indicating the correlation with the beginning of the event, but clearly shows that the main phase of the storm greatly outlasts the cosmic-ray increase. In fact, there is some evidence that the cosmic-ray increase occurs during the initial part of the main phase and that during the decay of the main phase the intensities have already returned to normal.

In a previous publication [Winckler and Bhavsar, 1960] we suggested that the alteration of Störmer cosmic-ray cutoffs might be produced by magnetic fields in space carried from the sun

during times of solar ejection of matter which join continuously to the dipole lines of force. Recent direct measurements show fields in space up to 50 γ completely free of the environment of the earth [Coleman, Sonett, Judge, and Smit, 1960]. Fields having the magnitude required to produce the observed changes in Störmer cutoff at the latitude of Minneapolis have not been observed and seem improbably large. Furthermore there does not seem to be a simple or plausible way in which these random space fields can be conceived to produce the systematic storm variations in the earth's surface field.

For many years the main phase of the magnetic storm has been associated with an equatorial ring current flowing in a westerly direction at a considerable distance from the earth. The existence of such ring currents exist in quiet times has been directly demonstrated [Smith, Coleman, Judge, and Sonett, 1960]. Ray [1958] has calculated the effects of such a permanent ring in altering the geomagnetic cutoffs for cosmic rays. Ray was concerned with applying this analysis to the 'knee' of the latitude effect in the normal cosmic-ray radiation and not to the case of the magnetospheric storm.

Kellogg and Winckler (private communication) have discussed the mechanics by which the ring current responsible for the main phase of the magnetic storm may lower the Störmer dipole cutoff energy, particularly at higher latitudes. Intensity changes may be caused by such a cutoff change operating on a fixed spectrum external to the earth by removing particles below a certain rigidity and leaving unchanged the intensity at each point in the spectrum above this rigidity. There is, however, no direct spectral evidence to support a change of this kind. On the contrary, nuclear emulsion measurements during the May 1959 event [Ney, Freier, and Winckler, 1959] show that simultaneous intensity changes of orders of magnitude occur in all parts of the spectrum in the measured region of direct observations between 80 and 250 Mev proton energy. A corresponding emulsion analysis is not yet available for the present event, nor is there a good way to get equivalent information concerning the presence or absence of a sharp Störmer-type cutoff from the counting instruments used on the flights discussed in this paper. But the presence of particles below the normal cutoff rigidity is easily identified

steep, intense spectrum frequently present. The above-mentioned association between lowering of cosmic-ray cutoffs, the development of a westerly flowing ring current, and the main phase of the storm is correct, then it can be expected that during the initial phase of the storm, when the horizontal surface field decreased, a corresponding increase in the geomagnetic cutoffs would occur. Measurements were made during the sudden commencements of all three of the July storms, as described in Table 1. No effect was observed in the first case

when only galactic cosmic rays were present. This may simply mean that no geomagnetically sensitive radiation was present in the primary cosmic rays at that time and latitude. On the second and third events a moderate intensity of solar cosmic rays was present during the sudden commencements as a precursor to the main storm increase. Although the presence of these particles before the main phase is not completely understood, nevertheless a drop in intensity was observed at the time of the sudden commencements in both the second and the

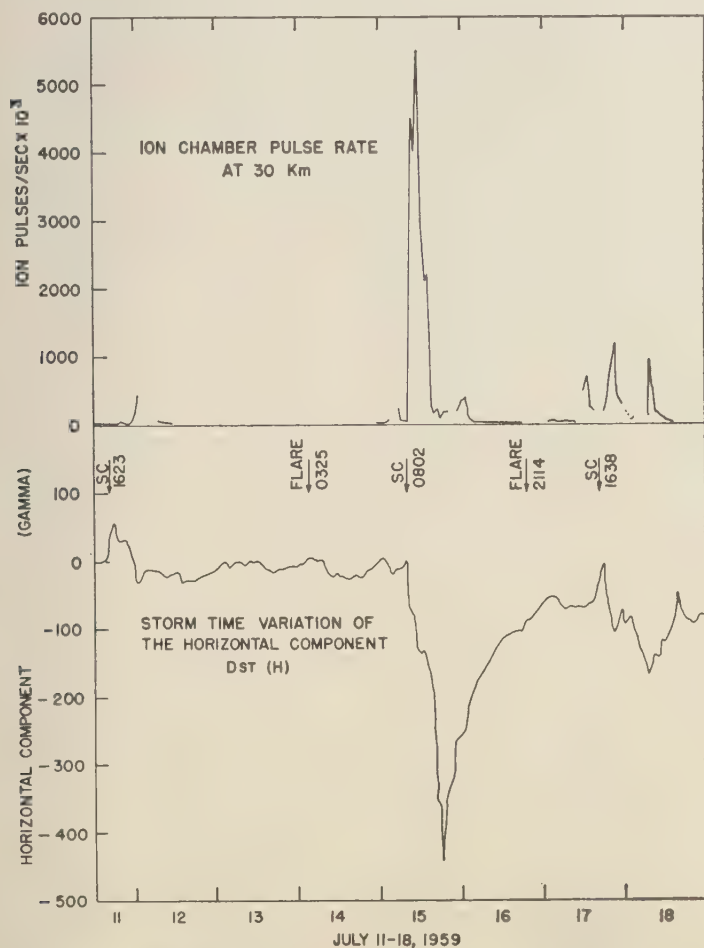


Fig. 22. Upper: Comparison of the high altitude ionization rates during all of the July events, and the $Dst(H)$ curve (Chapman and Akasofu) for the July events. Lower: The cosmic-ray increases are all shown on the same linear scale. The cosmic ray geomagnetic cutoffs at Minneapolis return to normal early in the main phase of the storm.

third storms. These field-sensitive particles apparently demonstrate an appreciable increase in the cutoff energy at the time of the sudden commencement, which would then be associated with an easterly current producing as well the positive phase of the storm. Such an easterly current flowing in the interface between the solar plasma cloud and the earth's field is necessary to account for the 'compression' of the surface field of the earth.

The free space intensity. It is of great interest to get all possible information about the full intensity of the solar cosmic ray flux in free space beyond the confines of the earth's magnetic field. This is of basic scientific interest with very important implications for the biological effects produced in space flight. The balloon measurements, even if made in the polar region, give direct values of the proton intensity only as low as about 80 Mev in the spectrum. Several indirect methods have been suggested for obtaining information about the low energy portion below the air cutoff at the balloon height. Freier (University of Minnesota, to be published) has examined the neutron intensity in emulsions during these events. These neutrons are produced in the atmosphere in nuclear collisions high above the balloon by the very intense low energy portion of the spectrum. Bhavsar (University of Minnesota, to be published) on many of these flights, as shown in the accompanying figures, has measured the intensity of photons that can be produced by nuclear collisions. In a way analogous with the neutron measurements, the γ -ray intensity from the nuclear collisions in the atmosphere high above the balloon, if properly interpreted, can give information about the low-energy portions of the spectrum. Bhavsar (to be published) estimates that during the May event [Winckler and Bhavsar, 1960] the cutoff in the spectrum at Minneapolis could have been as low as 40 Mev, assuming the exponent to remain unchanged down to this energy. At the latitude of Minneapolis it cannot, of course, be determined whether this cutoff is inherent in the spectrum in free space or whether it is a residual cutoff imposed by the magnetic field of the earth during the stormy conditions.

Estimates of the free space integral intensity can be obtained by assuming that the spectrum as measured in balloons can be extrapolated

TABLE 3. Free Space Flux from Spectrum Extrapolation

Energy Cutoff Mev	Total Flux Particles/cm ² ·sec·ster
88	2.4×10^2
40	2.4×10^3
20	1.8×10^4

down to any specific energy. The July 15 event is of great interest since this is the most intense one recorded to date. Using the spectrum measured at 1046 UT on July 15, it is possible to infer the integral intensity for a series of cutoffs as given in Table 3. We cannot, of course, be certain that the exponent remains fixed. There is evidence that it varies from one event to another, as shown by this paper and others describing these spectra (Ney, Winckler, and Freier, 1959; Freier, Ney, and Winckler, 1959).

A very unusual opportunity for comparing simultaneously the balloon-measured solar cosmic rays and the free-space cosmic rays was provided during the flight of the Pioneer V deep space probe launched March 11, 1960. A ionization chamber and counter were installed on the Pioneer V that were basically of the same nature as the balloon instruments used for the measurements described in this paper. A high level of solar activity occurring during late March and early April produced a series of cosmic-ray flares. On April 1, a 3+ flare occurred in which the solar cosmic-ray intensity was measured in space by Pioneer V and by a balloon flown at Minneapolis during the flare. The balloon was in the air at the time of the flare and also during a severe geomagnetic disturbance resulting from a previous flare in which, as described in this paper, the geomagnetic cutoffs were altered and the low-energy solar cosmic rays were observed again at Minneapolis. The details of the space measurements have been published [Arnoldy, Hoffman and Winckler, 1960b]. The details of the balloon measurement will be described shortly (Mallory, May, and Winckler, to be published). The balloon flight ascended during the solar proton event so that a range spectrum for the protons could be evaluated in the usual way. Both the balloon flight and the space probe clearly showed the decay of the protons in space, fi-

ing quite closely the T^{-2} law. The balloon spectrum was corrected for the decay of the particles in time, while the balloon ascended in order to evaluate the spectrum correctly. In the balloon measurements, valid over a range of proton energies of 130 to 425 Mev, the integral power-law spectrum exponent was found to be -2.4 , and this spectrum was then extrapolated and fitted the Pioneer V observations, provided a cutoff of about 30 Mev was used in space. The fit to the Pioneer V observations makes use of the ionization-per-count of the two instruments as well as the absolute intensities. These results provide some justification for the extrapolation of the spectrum as given in Table 3.

Knowledge of the spectrum shape is not necessary, the simultaneous measurement enables us to obtain an empirical value both for the ratio of the counter to the free space counter and for the balloon ion chamber to the free space ion chamber. This ratio empirically corrects for such factors as the limited solid angle in balloon measurements due to the steepness of the spectrum and the absorption of the particles incident at large zenith angles. The empirical ratio also includes the effects of the portion of the spectrum not seen by the balloon below the residual air cutoff. The ratios obtained

between the balloon flight and the space probe for the April 1 event have then been applied to the event occurring on July 15, seen only with balloons, to infer the free space intensities. A summary of all of these measurements is contained in Table 4. The free space intensity inferred at 1100 UT on July 15, 32 hours after the beginning of the flare, is calculated by comparing the Pioneer V and balloon flights. In estimating the peak intensity to be found in space during this event, the decay of the particles in space must be considered. For this event it is very difficult to get good information about the decay starting at the time of the flare. The event was so intense that many polar ionospheric records tended to be saturated and for the initial stages at least cannot be satisfactorily interpreted. If we assume that the decay law follows that seen clearly in space by Pioneer V in April 1960, with the variation with time proportional to T^{-2} , then we can extrapolate backward from the time of the balloon measurement to the time of the flare. One obtains the values in the last line of Table 4 of 1.5×10^7 cm²/sec for the omnidirectional particle intensity and of 7.6×10^3 R/hr for the dosage inside a typical instrument package with less than 1 g/cm² of stopping material. So far, there is absolutely no evidence that values as high as this are ever at-

TABLE 4. Examples of Galactic and Solar Cosmic Rays as Measured with Balloons and Space Probe

Type Experiment	Date	Time, UT	Particle Flux	Dosage Rate	Type Radiation
Space (Pioneer V)	1960 March 11-23	...	4.6/cm ² ·sec	0.0006 r/hr	Protons, α particles, etc. 10^8 - 10^{17} ev
Space (Pioneer V)	April 1	1200	33/cm ² ·sec excess	0.026 r/hr excess	Protons 40-500 Mev
Balloon	April 1	1200	0.4/cm ² ·sec excess	0.0004 r/hr excess	Protons 90-500 Mev
Balloon	2/cm ² ·sec	0.0004 r/hr	Galactic C.R.
Balloon	1959 July 15	1100	215/cm ² ·sec	0.140 r/hr	Protons 90-500 Mev
Pioneer V	April 1	1200	82 \times	65 \times	Protons
Balloon	July 15	1100	1.76×10^4 /cm ² ·sec	9.1 r/hr	Protons 40-500 Mev
Space (corrected)	July 14	0600	1.5×10^7 /cm ² ·sec*	7.6×10^3 r/hr*	Protons 40-500 Mev

These are extrapolated values and are highly uncertain (see text).

tained in interplanetary space. It is of course uncertain whether all events should be extrapolated back to the time of the flare using the T^{-2} law.

Acknowledgments. As in many times in the past, the Minnesota group was alerted that cosmic-ray flares were in progress by Mr. Harold Leinbach of the Geophysical Institute, University of Alaska, who observed these events on the riometer station at Fairbanks soon after the flares occurred. The regular flights during the IGC year, 1959, were sponsored by the Office of Naval Research. In addition, a grant from the National Science Foundation to implement a continuation of the high altitude monitoring program had been made and the program was underway at this time. The National Science Foundation project as a whole immediately participated to the fullest extent in flights made during this series. Mr. Ray Maas and Mr. William Huch provided assistance with the launchings, and Mr. James Stoddart contributed valuable experience in the rapid tracking and recovery of the equipment, which permitted the group to make additional flights with the quantity of apparatus at hand. We are indebted to Mrs. Gladys Seem for her painstaking work in the readout and plotting of the many strip charts accumulated from the telemetering and camera records for this series and for the entire IGY-IGC period as well. Members of the student group assisting with the preparation of equipment and launching of the flights included Ralph Fuchs, Dave Hofmann, Gerald Claessens, Pat Eckelberry, Thomas May, Andrew Masley, and many others.

This work was supported by the Office of Naval Research under contract Nonr-710(19) and by a grant from the National Science Foundation.

REFERENCES

- Anderson, K. A., and D. C. Enemark, Observations of solar cosmic rays near the north magnetic pole, *J. Geophys. Research*, **65**, 2657-2671, 1960.
- Arnoldy, R. L., R. A. Hoffman, and J. R. Winckler, Observations of the Van Allen radiation regions during August and September, 1959, I, *J. Geophys. Research*, **65**, 1361-1376, 1960a.
- Arnoldy, R. L., R. A. Hoffman, and J. R. Winckler, Solar cosmic rays and soft radiation observed at 5,000,000 kilometers from earth, *J. Geophys. Research*, **65**, 3004-3007, 1960b.
- Brown, R. R., and R. G. D'Arcy, Observations of solar flare radiation at high latitude during the period July 10-17, 1959, *Phys. Rev. Letters*, **3**, 390-392, 1959.
- Carmichael, H., and J. F. Steljas, Unusual cosmic-ray fluctuations on July 17 and 18, 1959, *Phys. Rev. Letters*, **3**, 392-394, 1959.
- Charakhchian, A. N., V. F. Tulinov, and T. N. Charakhchian, Cosmic rays emitted by the sun, p. 649, *Space Research, Proc. First International Space Science Symposium, Nice*, North-Holland Publishing Co., Amsterdam, 1195 pp., 1960.
- Coleman, P. J., Jr., C. P. Sonett, D. L. Judge, and E. J. Smith, Some preliminary results of the Pioneer V magnetometer experiment, *J. Geophys. Research*, **65**, 1856-1857, 1960.
- Fowler, P., P. S. Freier, and E. P. Ney, The primary alpha particle spectrum over North America and geomagnetic cutoff energies, *Nuovo cimento*, **8**, suppl., 492-499, 1958.
- Freier, P. S., E. P. Ney, and J. R. Winckler, Balloon observation of solar cosmic rays on March 26, 1958, *J. Geophys. Research*, **64**, 685-688, 1959.
- McDonald, F. B., Study of geomagnetic cutoff energies and temporal variation of the primary cosmic radiation, *Phys. Rev.*, **107**, 1386-1395, 1957.
- McDonald, F. B., and W. R. Webber, Proton Component of the primary cosmic radiation, *Phys. Rev.*, **115**, 194-205, 1959.
- National Bureau of Standards, Part B, *Solar Geophysical Data, CRPL-F180*, Central Radio Propagation Lab, Boulder, Colorado, August 1959.
- Ney, E. P., J. R. Winckler, and P. S. Freier, Protons from the sun on May 12, 1959, *Phys. Rev. Letters*, **3**, 183-185, 1959.
- Reid, G. C., and H. Leinbach, Low-energy cosmic-ray events associated with solar flares, *J. Geophys. Research*, **64**, 1801-1805, 1959.
- Smith, E. J., P. J. Coleman, Jr., D. L. Judge, and C. P. Sonett, Characteristics of the extraterrestrial current system: Explorer VI and Pioneer V, *J. Geophys. Research*, **65**, 1858-1861, 1960.
- Winckler, J. R., Balloon study of high-altitude radiations during the International Geophysical Year, *J. Geophys. Research*, **65**, 1331-1359, 1960.
- Winckler, J. R., and P. D. Bhavsar, Low-energy solar cosmic rays and the geomagnetic storm of May 12, 1959, *J. Geophys. Research*, **65**, 2637-2655, 1960.

(Manuscript received January 9, 1961.)

The High-Energy Cosmic-Ray Flare of May 4, 1960

1. High-Altitude Ionization and Counter Measurements

J. R. WINCKLER, A. J. MASLEY, AND T. C. MAY

*School of Physics, University of Minnesota
Minneapolis 14, Minnesota*

Abstract. Total ionization and counting rate measurements were made at 6 g/cm² depth at Minneapolis in the period from 7 to 16 hours following the cosmic-ray flare of May 4, 1960. The excess energy influx 7 hours after the flare was 340 Mev/cm²-sec in an atmospheric column. The omnidirectional ionization and counting rates were about 25 per cent above normal and the ionization ratio per particle was 1.2 times normal galactic cosmic rays at the same altitude. This is consistent with the high-energy nature of the event and contrasts greatly with the steep low-energy spectrum frequently observed during geomagnetic disturbances at Minneapolis. There is evidence for a displaced impact zone effect in the period 7 to 15 hours after the flare.

A balloon flight was made at Minneapolis carrying an integrating total ionization chamber, Geiger counters, and nuclear emulsions that measured cosmic rays originating in the flare of May 4, 1960. This flare occurred on the extreme east limb of the sun and was of 3+ magnitude; it began at 1015 UT and continued after 1105 UT. The flare was accompanied by radio noise bursts and noise storm continuum radiation over a wide frequency range, characteristic of cosmic-ray acceleration events. The balloon flight ascended at 1535 UT and reached its floating altitude at 1710 UT. The ionization rate plotted against pressure is shown in Figure 1 (flight M-59). For comparison, a flight during normal conditions (M-63) is used as a base for determining the excess incident energy. During the four ascent of flight M-59, the average energy flux in a column of the atmosphere brought by the flare particles and determined from the total ionization curves is 340 Mev/cm²-sec. This value, within the various errors, is consistent with that obtained from the emulsion analysis (see part 2). The excess ionization rate at 6 g/cm² atmospheric depth at approximately 1700 UT is about 25 per cent above the normal galactic cosmic-ray background ionization rate at that altitude. The vertical coincidence telescope carried in the monitoring unit likewise showed a similar increase above background at this time. If the omnidirectional ionization rate is

compared with the omnidirectional counting rate of one of the balloon Geiger counters, an ionization ratio only 1.2 times that of the normal galactic cosmic rays at the same altitude is found. This means that the primary particles producing the effects measured lie, on the average, in the high-energy range, contrasting greatly, for example, with the many cases of low-energy solar cosmic rays from flares reported recently [Winckler 1960; Winckler and Bhavsar, 1960; Ney, Winckler and Freier, 1959]. In the low-energy events, the ionization-per-count ratios were found to be 5 times or more the normal value at balloon altitude. If the normal cutoff at Minneapolis of about 300 Mev was in effect on May 4, and if the flare spectrum at that time was sufficiently flat, this effect could have been produced. Further discussion of this point is given in part 2 of this paper.

The decay of the event observed at balloon height is shown in Figure 2 between 1700 UT on May 4 and 0200 UT on May 5 when the balloon flight terminated. Excess radiation was seen throughout this flight but with decreasing intensity. There is no clear evidence for a change in the nature of the radiation during this time. Figure 2 also shows the counting rate profile of the University of Minnesota neutron monitor that measured the initial part of the event during the flare. This monitor gave useful information, although the statistical accuracy is not

particularly high. It showed that the first increase occurred close to 1030 UT on May 4 and that the sea-level effects had essentially disappeared by 1145 UT. In Figure 2, the balloon telescope and ion chamber base rates have been taken from a balloon flight on May 7. The neutron monitor record showed steady conditions between the end of the flare event and May 7. This at least assures that the high-energy galactic background was the same, so that the high-altitude base rate may be determined with a fair degree of certainty.

In Figure 3 we compare the Minnesota neutron monitor flare radiation and the balloon

measurement of the flare radiation plotted against hours from the flare (assumed to be 1024 UT) on a log-log scale. The neutron monitor points can be fitted with a slope showing that the time decay is proportional to T^{-2} , although this exponent is not very precisely determined by the measurement. The ionization excess and the vertical coincidence excess at the time of the balloon flight show no clear evidence for a definite slope. The rates at first drop rapidly, then show a minor plateau up to 10 hours after the flare, and then the radiation rapidly disappears with a very steep time dependence. Unfortunately, the balloon was not in the air dur-

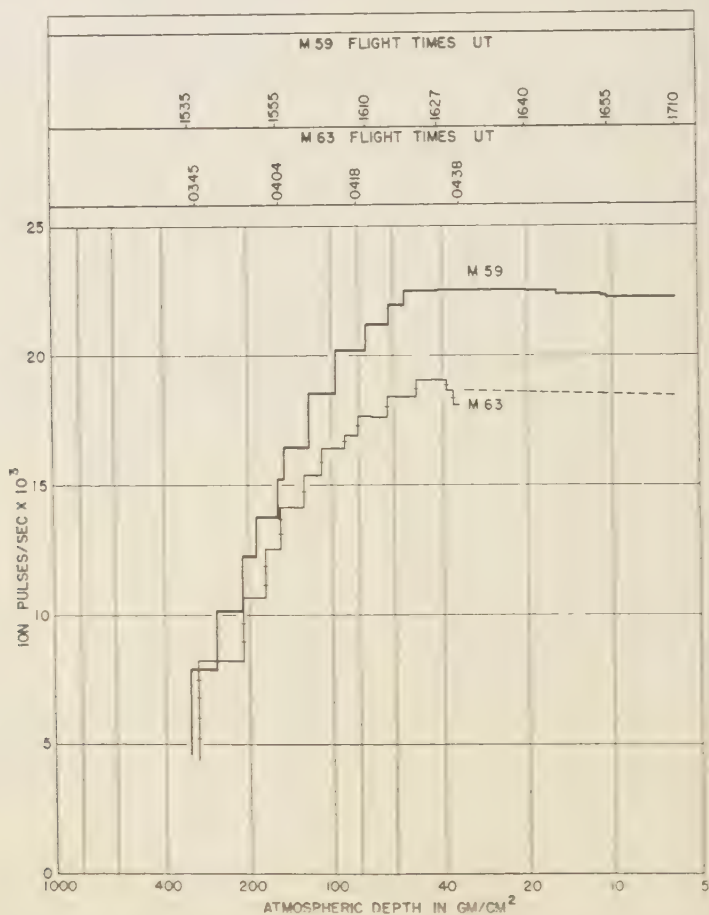


Fig. 1. Normalized ionization-depth curves for the solar flare radiation after 7 hours (M-59), and a normal day sounding (M-63). The normalization is the same as that plotted for the IGY monitoring experiments [Winckler, 1960].

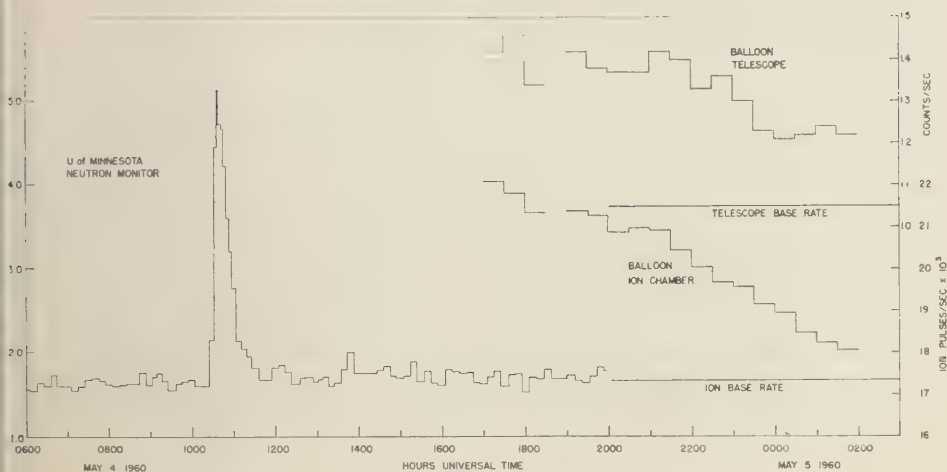


Fig. 2. Flare observations at Minneapolis at sea level and high altitude. The vertical coincidence telescope and balloon ion chamber data are taken from the constant level portion of the monitoring flight.

g the time that the neutron monitor showed the initial flare radiation; there was not time after the flare alarm to get the balloon to high altitude since the sea level effect lasted for about an hour only. We have, however, extrapolated back the high-altitude radiation, assuming a slope -2 like that of the neutron monitor. Alternatively, the extrapolation of the neutron monitor to high altitude may be accomplished using the ratio of balloon changes to neutron monitor changes during Forbush-type intensity decreases in 1956 when low-energy radiation was present in the galactic primaries [Winckler, 1960]. This procedure results in a slope of -1.4 . Whichever curve is taken, as shown in Figure 3, the high-altitude radiation at the beginning of the event must have exceeded normal rates by at least a factor of 100, perhaps by 500 or more. The balloon measurements are consistent with the assumption that the flare radiation decayed from flare onset to the time of the balloon flight with a decay law of $T^{-1.5}$ or T^{-2} . These decay estimates are admittedly crude and are mainly intended to show that the flare cosmic rays decreased in a manner similar to other such events [Meyer, Parker and Simpson, 1956; Arnoldy, Hoffman, and Winckler, 1960].

The high altitude intensity from 7 to 15 hours after the flare shows some evidence for the passage of the longitude of Minneapolis through an

impact zone. However, the time of day is completely wrong for impact zones which have been studied by Firor [1954]. If the rapid decrease of the radiation at 0100 UT on May 5 is an impact zone effect, then the local time dependence must be shifted in phase from that given by Firor by 4 hours. This may not be impossible in view of recent evidence for semipermanent ring currents caused by trapped radiation at large distances from the earth [Smith, Coleman, Judge, and Sonett, 1960]. A preliminary examination of sea-level neutron monitors shows, however, that striking impact zone effects occurred during the first hour of the event. The cosmic rays were seen over North America particularly and not in Europe or Japan.

A further balloon flight after M-59 showed at most a 5 per cent excess that may be attributed to flare radiation on the following day, May 5, from 0400 to 0800 UT. There was no ensuing geomagnetic storm directly related to this flare. In some respects, the May 4 flare resembled the February 23, 1956, event in that it contained mainly high-energy radiation, was located on the west limb of the sun, and was notable for the rapid arrival of the cosmic ray particles from the flare itself. The relatively low intensity of low-energy protons is also shown by the fact that the polar cap ionospheric effects as determined by riometer measurements were not large

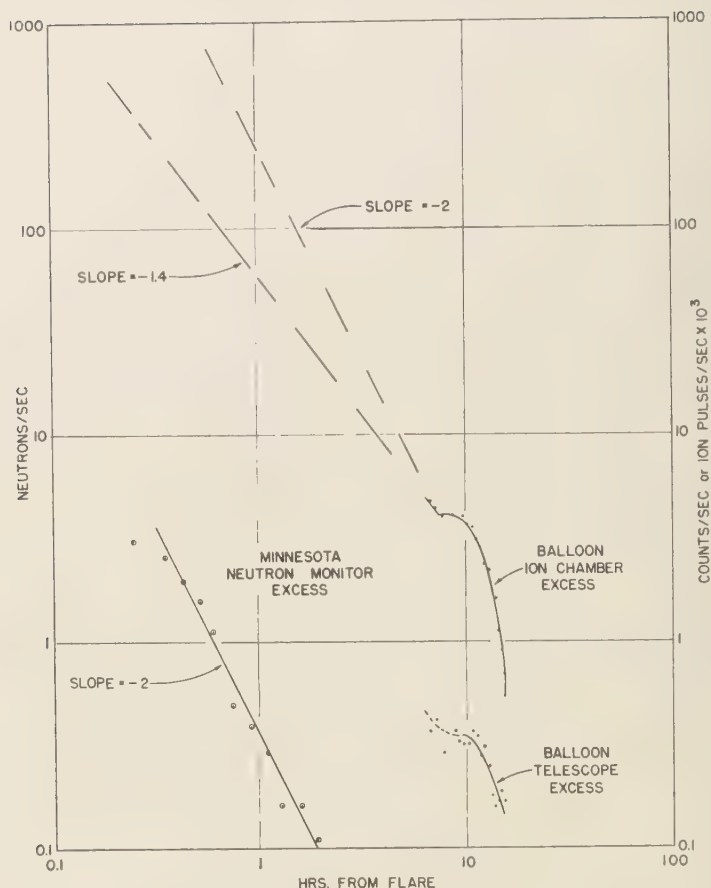


Fig. 3. Decay of the excess radiation seen at sea level and at high altitude.

[Leinbach, 1960; Kiruna, 1960]. If the spectrum had extended to low energies and had had the spectral shape that has been measured for the low-energy cosmic-ray events, huge polar-cap disturbances would have been produced. There are two alternatives: (a) that the flare mechanism itself placed the cosmic radiation into the high-energy region alone, or (b) that the low-energy particle content of this flare could not reach the earth because of the magnetic conditions in space. Evidence against assumption (b) is that in a number of cases the low-energy type of solar cosmic rays has been observed from flares in the western quadrant of the sun at large longitudes [Freier, Ney, and Winckler, 1959; Winckler, 1960]. The solar longitude effect has been discussed recently by

several authors, but no very decisive picture has as yet emerged [McCracken and Palmeir, 1960].

Acknowledgments. This work was supported by a grant from the National Science Foundation. We are grateful to Professor John Simpson of the University of Chicago who roused us in the early morning hours with the cosmic-ray flare alarm. Assistance with this experiment was provided by members of the University of Minnesota high altitude monitoring group.

REFERENCES

- Arnoldy, R. L., R. A. Hoffman, and J. R. Winckler, Solar cosmic rays and soft radiation observed at 5,000,000 kilometers from earth, *Geophys. Research*, 65, 3004-3007, 1960.
- Firor, John, Cosmic radiation intensity-time variations, *ibid.*, 65, 3008-3010, 1960.

- ons and their origin, IV, Increases associated with solar flares, *Phys. Rev.*, *94*, 1017-1028, 1954.
- er, P. S., E. P. Ney, and J. R. Winckler, Balloon observation of solar cosmic rays on March 3, 1958, *J. Geophys. Research*, *64*, 685-688, 1959.
- una *Geophysical Data*, Data summary no. 5, Kiruna Geophysical Observatory, Royal Swedish Academy of Science, Kiruna, May 1960.
- abach, Harold, The polar-cap absorption events of March 31 through May 13, 1960, Technical report, Geophysical Institute, University of Alaska, College, Alaska, June 21, 1960.
- yer, P., E. N. Parker, and J. A. Simpson, Solar cosmic rays of February 1956 and their propagation through interplanetary space, *Phys. Rev.*, *94*, 768-783, 1956.
- Cracken, K. G., and R. A. R. Palmeira, Comparison of solar cosmic ray injections including July 17, 1959, and May 4, 1960, *J. Geophys. Research*, *65*, 2673-2683, 1960.
- Ney, E. P., J. R. Winckler, and P. S. Freier, Protons from the sun on May 12, 1959, *Phys. Rev. Letters*, *3*, 183-185, 1959.
- Smith, E. J., P. J. Coleman, Jr., D. L. Judge, and C. P. Sonett, Characteristics of the extraterrestrial current system: Explorer VI and Pioneer V, *J. Geophys. Research*, *65*, 1858-1861, 1960.
- Winckler, J. R., Balloon study of high-altitude radiations during the International Geophysical Year, *J. Geophys. Research*, *65*, 1331-1359, 1960.
- Winckler, J. R., and P. D. Bhavsar, Low-energy solar cosmic rays and the geomagnetic storm of May 12, 1959, *J. Geophys. Research*, *65*, 2637-2655, 1960.

(Manuscript received January 16, 1961.)

The High-Energy Cosmic-Ray Flare of May 4, 1960

2. Emulsion Measurements

S. BISWAS¹ AND P. S. FREIER

*School of Physics, University of Minnesota
Minneapolis 14, Minnesota*

Abstract. The differential energy spectrum of solar protons from the flare of May 4, 1960, has been measured in emulsions during a balloon flight from 1700 UT on May 4 to 0200 UT on May 5. An increased flux of protons of rigidity 0.7 to 1.6 bv of 600 ± 150 particles/cm²·sec·ster was measured. The differential rigidity spectrum in this interval can be expressed as

$$\frac{dN}{dR} = \frac{(0.65 \pm 0.15) \times 10^{-3}}{R^{1.1 \pm 0.5}} \text{ protons/m}^2 \cdot \text{sec} \cdot \text{ster} \cdot \text{bv}$$

There were no solar α particles present during this time. The flux of $Z \geq 3$ nuclei was normal.

The emulsion measurements. In an attempt to identify and measure the particles responsible for the increased counter and chamber rates during the flare event of May 4, 1960 (as discussed in part 1), we have measured, in the nuclear emulsions, the flux of particles in four different types of scanning:

1. A scan for particles of ionization $\geq 9 I_{\min}$. This gives us the flux of nuclei of $Z \geq 3$. Corrected to the top of the atmosphere this flux is

$$F^0(Z \geq 3) = 12.8$$

$$\pm 2.2 \text{ particles/m}^2 \cdot \text{sec} \cdot \text{ster}$$

This is a normal flux value for this time, when the galactic cosmic rays were recovering from a sharp decrease that occurred on April 30. From this scan we also obtained the differential flux of particles of energy 100–240 Mev/nucleon at the top of the atmosphere. Figure 1 shows the differential flux of these particles on May 4 compared with the normal galactic flux that is obtained during years of sunspot maximum but with no occurrence of a solar flare. The differential flux of solar α particles expected during the flight is also shown. This spectrum was calculated from the measured rigidity spectrum of the solar protons, assuming that α particles are 15 per cent as abundant as protons of the same rigidity. There is no evidence for solar α particles from

the flare of May 4, 1960. This is in contradiction to the observations of solar α particles from the flares of May 10, 1959, July 10, 1959, and September 3, 1960 (University of Minnesota emulsion results, to be published). A value of the differential flux of protons from 40 to 57 Mev at the top of the emulsion stack was also obtained. These data are plotted in Figure 2 and will be discussed later. These protons are entirely secondaries.

2. A scan for particles whose ionization is $\geq 3 I_{\min}$. From this scan the total flux of α particles was obtained. Energies of the particles were determined by scattering, and for energies ≤ 600 Mev/nucleon we can determine the energy to ± 20 per cent by this method. For particles with energy < 240 Mev/nucleon the energy was determined by range. The integral α -particle flux is shown in Figure 3. Also shown is the flux obtained on June 14, 1958, near the maximum of the solar cycle but during a time when the sun had shown little activity. The flux on May 4, 1960, looks entirely normal and we cannot account for the increased rates with α particles. We also obtained from this scan the differential flux of protons in the interval 55–125 Mev at the top of the emulsion. The data are plotted in Figure 2.

3. A scan for particles of all ionizations. Since so far we had seen no abnormally high flux in α particles or heavy nuclei, we scanned the emulsions under a magnification of 15×100

¹ On leave from the Tata Institute of Fundamental Research, Bombay.

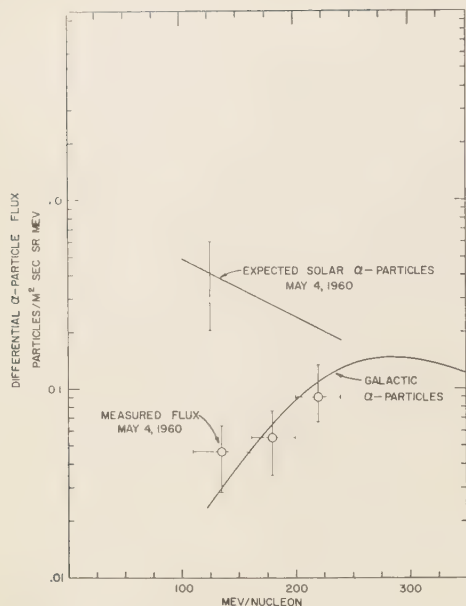


Fig. 1. The differential energy spectrum of α particles on May 4, 1960. The measured flux agrees with the galactic spectrum measured during 1958. The expected solar α particle flux has been calculated from the measured rigidity spectrum of solar protons assuming that α particles are 15 per cent as abundant as protons of the same rigidity.

and made ionization and scattering measurements on all particles passing through the emulsion. Although only 108 particles in the flight plates and 31 particles in an equal area of background plates were measured by this method, we can obtain at least a rough idea of the total flux passing through the emulsion. After correcting for background tracks (this correction is for particles made before the balloon ascent), we obtained the following vertical fluxes:

electrons 15–150 Mev...	300 ± 150
	particles/m ² ·sec·ster
protons 200–600 Mev...	620 ± 170
	particles/m ² ·sec·ster
protons > 1.1 bev...	250 ± 500
	particles/m ² ·sec·ster

The background correction to the > 1.1 bev protons was large (a result of sea-level mesons going through the emulsions from the time of manufacture to development); hence the flux of high energy protons has a large uncertainty. The

only indication of an excess flux is the flux of protons 200–600 Mev. (The background correction is small for these particles.) We decided to measure the flux of these protons with increased statistics.

4. A scan for particles of ionization $\geq 1.1 I_{\min}$ and zenith angle $< 26^\circ$. Protons whose kinetic energy at the top of the atmosphere is 900 Mev have ionization of $1.13 I_{\min}$ at the scan line in the emulsion. A scan was made for all particles and a careful ionization measurement with a standard deviation of 7 per cent was made on each particle. An ionization distribution of all particles found in this scan was plotted, but only those particles whose ionization was 10 per cent above minimum were measured further. From a total of 326 particles found on this scan, 6 particles whose initial ionization measurement was $> 1.1 I_{\min}$ were obtained. Scattering and ionization measurements were again made on these particles. Protons of energy < 900 Mev are completely resolved from other particles with the exception of electrons of energy 790–1580 Mev whose ionization is at the plateau value of ionization rather than I_{\min} . In emulsion the plateau due to the relativistic ionization increase is ~ 10 per cent. To resolve these electrons, α particles in this interval were followed a distance ≥ 3 cm and scattering measurements were made again. (Radiation length = 2.9 cm in emulsion.) The electrons will radiate in this distance and reveal their identity by their decreased energy. In this flight, however, only one electron was found in this group of 19 particles; the other 18 particles were protons. The differential proton flux obtained by this scan is plotted in Figure 2.

The proton differential energy spectrum. Figure 2 plots the differential proton flux values obtained in the various scans described. We are measuring protons of three kinds: (1) galactic primary protons whose differential spectra are shown in Figure 2 [McDonald, 1959]; (2) secondary protons from interactions in the air above the detector; and (3) solar protons resulting from the flare. The secondary proton spectrum has been measured in several flights similar in altitude to this flight. The differential flux of these secondary protons is shown in Figure 2. This spectrum for secondary protons is consistent with the secondary proton flux measured by McDonald and Webber [1959], and the energy spectra of secondaries measured in emulsion by various

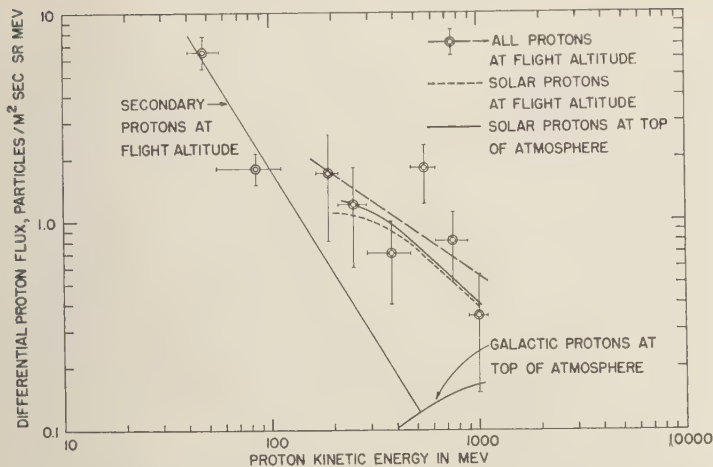


Fig. 2. The average differential energy spectrum of solar protons from 1700 UT on May 4 to 0200 UT on May 5. From the measured spectrum at flight altitude both the secondary and galactic protons have been subtracted to obtain the solar protons at flight altitude. From this the spectrum at the top of the atmosphere has been calculated.

investigators [Powell, Fowler, and Perkins, 1959]. The differential proton flux values measured during this flight are shown in Figure 2. The measured flux is significantly higher than both the secondary and galactic flux. After subtracting the secondary and galactic we obtained the differential energy spectrum of the remaining protons attributed to the solar flare. We have plotted the differential energy spectrum of these solar protons at the top of the atmosphere. The spectrum can be represented by

$$\frac{dN}{dE} = \text{const} \cdot E^{-(1.0 \pm 0.3)}$$

from 250–1000 Mev. The differential rigidity spectrum can be expressed as

$$\frac{dN}{dR} = \frac{(0.65 \pm 0.15) \times 10^{-3}}{R^{1.1 \pm 0.5}}$$

particles/m²·sec·ster·bv

for rigidity $R = 0.7 - 1.6$ bv.² This solar proton spectrum is much flatter than the measured spectra of solar protons from the May 10, July

10, July 14, and July 16, 1959, flares (University of Minnesota emulsion results, to be published).

Discussion of results. We have measured the average differential spectrum of solar protons from 1700 UT on May 4 to 0200 UT on May 5 in the rigidity interval 0.7 to 1.6 bv and found that

$$\frac{dN}{dR} = \frac{(0.65 \pm 0.15) \times 10^{-3}}{R^{1.1 \pm 0.5}}$$

particles/m²·sec·ster·bv

This spectrum must become as steep as R^{-7} for $R > 2$ bv in order to agree with the response of the sea-level neutron monitor and meson telescopes. If the spectrum had not become steeper, these sea-level detectors would have shown a large effect at the flight time; such an effect was not observed.

The measurements described in part 1 indicate that the solar particles were decaying with time as $\sim T^{-2}$ after 1040 UT. Assuming that particles of all rigidities decay at this rate, we can derive the rigidity spectrum at 1040 UT from the spectrum during the flight time. To fit the spectrum most smoothly over the entire rigidity region we have used the steeper limit of $R^{-1.5}$ on our measured spectrum and extrapolated to higher rigidities with an R^{-7} spectrum. This spectrum is shown in Figure 4 for the average

² No primary particles of rigidity < 0.7 bv were observed although primary particles of rigidity > 1.45 bv could reach the detector. Because of the low solar proton flux and high secondary flux correction, it is difficult to measure accurately the cutoff rigidity in this flight.

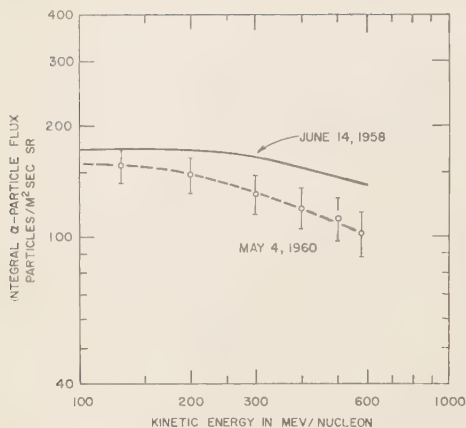


Fig. 3. The integral α particle flux measured from 1700 May 4 to 0200 UT May 5. The May 4 flux is lower than the June 14, 1958, flux. On May 4 the galactic cosmic rays were recovering from a Forbush decrease which occurred on April 30.

flight time. Using the spectrum with a T^{-2} time dependence and $T = 0$ at flare onset, we have calculated the response of sea level neutron monitor³ and meson telescope [McCracken and Palmeira, 1960] using specific yield functions of Webber and Quenby [1959]. At 1040 UT when these sea-level detectors recorded maximum intensities the calculated increase is 240 per cent compared with a measured increase of 220 per cent for the Deep River neutron monitor. At the same time the calculated increase for the M. I. T. meson telescope is 3 per cent compared with a measured 5 per cent increase. At later times the agreement is also satisfactory. Therefore, we see that the derived rigidity spectrum fits both the flight data and the response of the two sea level detectors. If the spectrum is extrapolated to lower rigidities with a slope -1.6 , it would predict a riometer absorption of about 4 db [Webber, 1961] for cutoff rigidity of 0.46 bv at 1040 UT. The riometer at Kiruna (cutoff $\approx .47$ bv) recorded a peak absorption of 3.7 db at 1130 UT (J. Ortner, private communication), and that at Thule about 1 db at 1130 UT [Webber, 1961].

At 1040 UT the integral flux of solar protons

³ IGY Data Center A, Summary of May 4 data. We are grateful to Dr. H. Carmichael for supplying us the May 4 one-minute data of Deep River neutron monitor.

of rigidity ≥ 0.7 bv calculated from this spectrum is 70 particles/cm²·sec·ster. This is about 700 times the normal cosmic-ray flux, which is in agreement with the conclusions drawn in part 1. During the period of our flight the average integral flux of solar protons was about 600 particles/m²·sec·ster of rigidity ≥ 0.7 bv. This is about 0.6 times the normal cosmic-ray flux.

Using the rigidity spectrum during the flight period (Fig. 4) and the time decay as T^{-2} , we have calculated the total excess energy from solar protons during the period of the ascent of the balloon. We obtained the energy flux of 490 Mev/cm²·sec at the top of the atmosphere during the ascent. The measurements in part 1 showed that total energy loss in the column of atmosphere during the ascent of the balloon was 340 Mev/cm²·sec, which is consistent with our results within the uncertainties of the calculations and measurements.

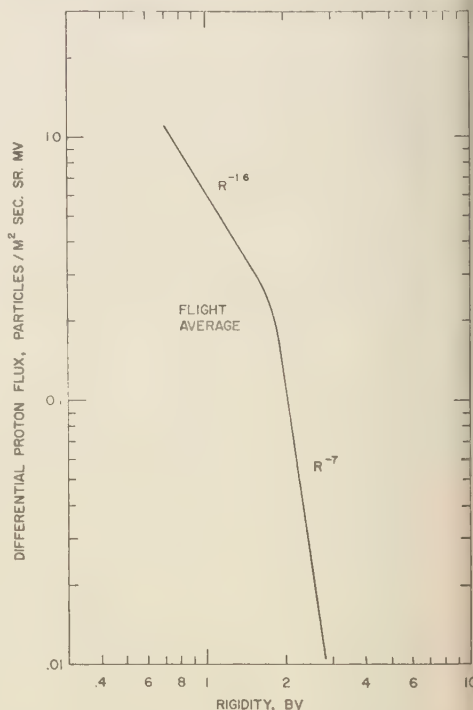


Fig. 4. The average differential rigidity spectrum of solar protons from 1700 UT May 4 to 0200 UT May 5. From 0.7 to 1.6 bv the spectrum is measured; for higher rigidity, the spectrum has been derived.

In conclusion we may summarize that the differential rigidity spectrum of solar protons had a slope of about -1.5 in the range 0.7 to 1.6 during the flight period. In the lower rigidity range the spectrum is most probably flatter and in the high rigidity range steeper with a slope of about -7 . The flare of July 10, 1959, had a measured proton differential spectrum of slope about -5 in the rigidity range 0.4 to 0.7 bv and a flat spectrum in the region of 0.3 to 0.4 bv (University of Minnesota emulsion results, to be published.) The flares we have studied so far indicate that many of the flares have similar differential rigidity spectra, being flat at some low rigidity and becoming steeper at higher rigidities. The flares are different in that the position of the flat part of the spectrum occurs at different rigidities and in some flares all particles observed at earth are in the steep part of the spectrum.

Acknowledgments. This work was supported by the joint program of the Office of Naval Research and the U. S. Atomic Energy Commission under Contract Nonr-710(19) and by a grant from the National Science Foundation. We wish to thank the monitoring group at Minnesota, who so faith-

fully fly balloons and never forget the can of emulsions. We also thank Miss Barbara Berry, Miss Julia Petts, Mrs. Hilda Rahlenbach, Mrs. Dawn Copeland, and Mrs. Biruta Sommer for their always painstaking and sometimes very accurate scanning of the emulsions.

REFERENCE

- McCracken, K. G., and R. A. R. Palmeira, Comparison of solar cosmic ray injections, including July 17, 1959, and May 4, 1960, *J. Geophys. Research*, **65**, 2673-2683, 1960.
- McDonald, F. B., Primary cosmic-ray intensity near solar maximum, *Phys. Rev.*, **116**, 462, 1959.
- McDonald, F. B., and W. R. Webber, Proton component of the primary cosmic radiation, *Phys. Rev.*, **115**, 194-205, 1959.
- Powell, C. F., P. H. Fowler, and D. H. Perkins, *The Study of Elementary Particles by the Photographic Method*, Pergamon Press, London, 669 pp., 1959.
- Webber, W. R., *Progress in Cosmic Ray Physics*, **VI**, North Holland Publishing Co., Amsterdam, to be published, 1961.
- Webber, W. R., and J. J. Quenby, On the derivation of cosmic ray specific yield functions, *Phil. Mag.*, **4**, 654-664, 1959.

(Manuscript received January 16, 1961.)

Large-Scale Electron Bombardment of the Atmosphere at the Sudden Commencement of a Geomagnetic Storm

R. R. BROWN

*Department of Physics, University of California
Berkeley, California*

T. R. HARTZ

*Defence Research Telecommunications Establishment, Shirley Bay, Ottawa
Ontario, Canada*

B. LANDMARK

*Norwegian Defence Research Establishment
Kjeller, Lillestrøm, Norway*

H. LEINBACH

*Geophysical Institute, University of Alaska
College, Alaska*

AND

J. ORTNER

*Kiruna Geophysical Observatory
Kiruna, Sweden*

Abstract. A burst of X rays was observed at balloon altitude over Alaska with the onset of a sudden commencement geomagnetic storm at 0146 UT on June 27, 1960. The electron bombardment of the upper atmosphere that gave rise to the X rays occurred on a large scale, ionospheric absorption coincident with the X-ray burst being observed by riometers in Alaska, Sweden, and Norway.

Introduction. The occurrence of X-ray effects at high altitude in connection with auroral and magnetic disturbances [Van Allen, 1957; Winckler, Peterson, Arnoldy, and Hoffman, 1958; Winckler, 1960; Anderson, 1958] has been a problem of geophysical interest in recent years. Extensive observations, using balloon and rocket techniques, have demonstrated that these effects arise from the incidence of low energy electrons on the top of the atmosphere. The origin of these particles and the location of the region where they acquire their kinetic energy have been discussed extensively in the literature.

Anderson [1958], observing a close correlation between the presence of these X rays and a geomagnetic storm, has suggested that the electrons responsible for the X rays are local in character, i.e., either they are accelerated in the

vicinity of the earth or they acquire their energy near the sun and are released locally when the solar plasma cloud bearing them encounters the geomagnetic field.

It is the purpose of the present paper to report observations of an electron bombardment of the atmosphere that was in time-coincidence with the onset of a sudden commencement geomagnetic storm. The electrons, by their ionizing effects in the upper atmosphere, produced a brief increase in ionospheric absorption of cosmic radio noise, observed by a number of riometers, [Little and Leinbach, 1959] as well as X-ray effects, by bremsstrahlung, which were observed by a balloon-borne radiation instrument. These observations serve to amplify the type of event reported by Anderson [1958] by indicating the ionospheric absorption at the time of the X-ray burst and the geographical

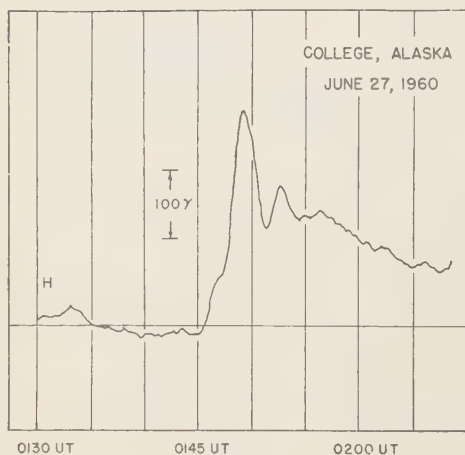


Fig. 1. Rapid-run magnetogram of the horizontal component H of the geomagnetic field (College, Alaska).

extent of the absorption event. In addition, estimates of the energy and flux of the electrons were obtained from the balloon observations.

Solar and magnetic activity. A sudden commencement geomagnetic storm began at 0146 UT on June 27, 1960. This storm is thought to have resulted from a flare of importance 3 reported some 29 hours earlier, starting at 2042 UT on June 25, 1960. At College, Alaska, the sudden commencement of the storm involved an increase of the horizontal component H of the geomagnetic field by about 330γ in 3.5 minutes and an increase in the vertical component Z of about 65γ in 2 minutes. A second maximum of the horizontal component H occurred at 0152–0153 UT. A reproduction from a rapid-run magnetogram of the sudden commencement phase of the increase in H is shown in Figure 1. The initial peak value of H was reached at 0149 UT, the rate of change between the onset of the sudden commencement and the maximum value of H amounting to about $90 \gamma/\text{min}$.

Following the sudden commencement, the magnetic storm, as shown by variations in H at College, involved three phases. The first, a positive-going phase, occurred between 0156 UT and about 0600 UT; during this time H fluctuated more slowly than during the initial phase and reached a maximum excursion of about 650γ above the prestorm level. This phase was fol-

lowed by a second, rather steady, phase during which changes in H were small and not particularly rapid. Then, starting at about 0845 UT the third, negative-going, phase developed and the maximum excursion in this direction was about 870γ below the prestorm level. This last phase ended at about 1100 UT and the H component returned toward normal.

Balloon observations. At the time of the magnetic storm a balloon-borne radiation instrument was aloft over central Alaska and a burst of radiation, presumably X rays, was observed in time-coincidence with the sudden commencement of the geomagnetic storm. The balloon flight was launched at 1615 UT, June 26, 1960, from Happy, Alaska (near College). Following the initial ascent, the balloon floated at a pressure altitude of 12 mb, drifting westward at about 25 knots. At the time of the magnetic storm, the balloon was between 250 and 275 km west of the launch site. The radiation detectors included in the balloon instrumentation consisted of two Geiger counters—one with a 30 mg/cm^2 cathode of Al and the other with a copper screen cathode plated with 100 mg/cm^2 of Bi.

With the onset of the sudden commencement geomagnetic storm at 0146 UT, the counting rate of both detectors increased. The counting rate of the Al counter increased by about 80 per cent, reaching a peak at 0149 UT and then returning to normal by about 0151 UT; the rate of the counter with the Bi plated cathode increased by about 70 per cent and followed the time variation of the Al counter rather closely. A second increase in the counting rates of the detectors occurred, this time about 1/2 as large as before, and after reaching a peak at about 0152 UT returned to normal at 0156 UT. These features of the radiation intensity variations are shown in Figure 2.

The intensity increase described above represented the most striking feature of this particular balloon flight. It was preceded, however, by two small increases of about 10 per cent amplitude, presumably of auroral origin, which lasted about 60 and 35 minutes, respectively. Following the burst of radiation starting at 0146 UT, a slow decrease in counting rate set in as a Forbush decrease of cosmic radiation developed.

On the assumption that the radiation burst

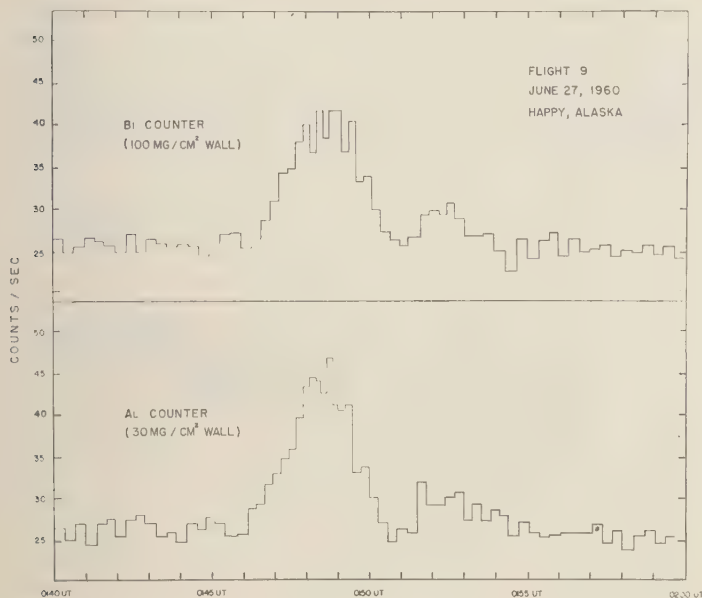


Fig. 2. Counting rate of balloon-borne detectors as a function of time.

sisted of X rays, as has been found to be the case during auroral and magnetic bay disturbances [Winckler, 1960], the ratio of the counting rates of the balloon-borne detectors indicates that the energy of the radiation lies in the region near 50 kev.

Riometer observations. Increased ionospheric absorption produced by the electrons responsible for the burst of X rays was observed with two riometers located at College, Alaska. Both riometers were in operation at 27.6 Mc/s but one had its antenna directed vertically upward while the antenna of the other was oriented about 15° east of north and placed so that the main lobe of the antenna was most sensitive to cosmic radio noise coming obliquely through the ionosphere at an elevation angle of about 30°. The vertical riometer recorded a sharp increase in absorption beginning at the onset of the sudden commencement storm. This absorption reached a maximum value of approximately 3.5 db at 0149 UT, while the oblique riometer recorded a maximum of about 5 db at this time. The ratio of the absorption recorded on the two beams is consistent with the hypothesis of a uniform absorbing layer extending over a north-south distance of at least 100 km.

Riometers at Kenai and Skwentna, Alaska, were in operation at the time of this event also (Little, private communication). The 50 Mc/s instrument at Kenai was unfortunately undergoing calibration at the time of the sudden commencement. At Skwentna, however, both the 30 Mc/s and the 50 Mc/s instruments recorded increased absorption beginning at about 0146 UT. The increased absorption of 30 Mc/s amounted to 1.0 db; on 50 Mc/s, the increased absorption was smaller, as would be expected.

Similar observations of increased ionospheric absorption in time-coincidence with the sudden commencement were obtained with riometers located at Kiruna, Sweden, and Alta and Harstad, Norway, almost directly across the polar cap from the Alaskan riometer sites. The Kiruna riometer, also operating on 27.6 Mc/s, recorded an increase in absorption amounting to 1.3 db while the riometers at Alta and Harstad showed 0.8 and 0.6 db of attenuation on 28.6 Mc/s, respectively.

These observations are represented, in part, in Figure 3 which includes the riometer records from College and Kiruna. These records display cosmic radio noise, with solar noise bursts superimposed, as a function of time. The signal levels

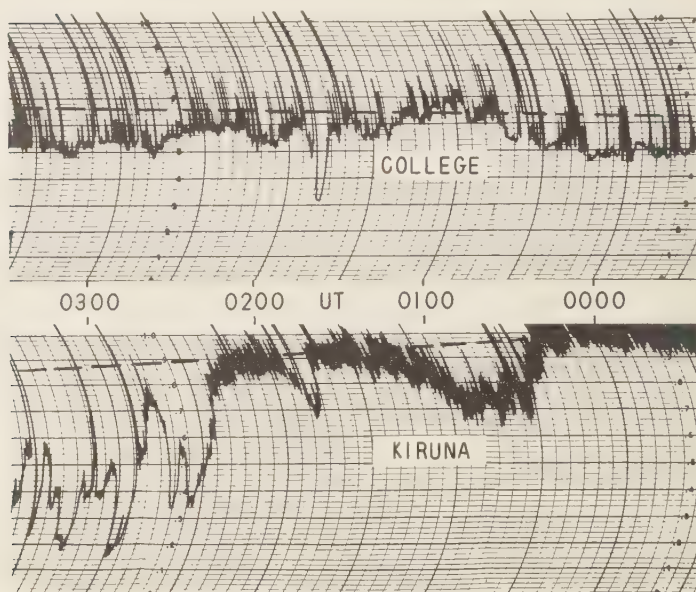


Fig. 3. Riometer records from College, Alaska, and Kiruna, Sweden, June 27, 1960. The signal levels expected under quiet ionospheric conditions are shown by the dashed lines.

expected under quiet ionospheric conditions are shown by the dashed lines; increased absorption is indicated by a reduction of the recorder current below the quiet level. Here, it is seen that both riometers were recording auroral absorption before and after the event, with stronger absorption at Kiruna, but the increase at the time of sudden commencement was the only absorption having simultaneity at the two stations.

In addition to the riometers that definitely recorded increased ionospheric absorption with the onset of the sudden commencement, seven other riometers were in operation at the time of the event. All the riometers are listed, in order of decreasing geomagnetic latitude, in Table 1. In addition to the geomagnetic latitudes of the riometers, their geographic coordinates are given in Table 1 as well as the elevation angle of the sun above the horizon at the time of the sudden commencement.

With the exception of Cape Jones, Canada, there is no evidence of increased absorption at the time of the event at the other riometer sites. The Cape Jones record, however, showed strong auroral absorption before, during, and after the event, and it is impossible to distin-

guish unambiguously a small increase in absorption at 0146 UT from variations that might be attributed to the auroral absorption in progress at the time.

Interpretation. The ionospheric absorption accompanying the sudden commencement of the geomagnetic storm is interpreted as resulting from ionization by a brief influx of electrons, possibly representing a 'dumping' of electrons from the outer radiation zone brought about by the rapid geomagnetic field variations at the time of the sudden commencement or the escape of electrons from plasma clouds encountering the geomagnetic field.

Another interpretation which could account for the ionospheric absorption observed by the riometers in Alaska, Sweden, and Norway is that the bombardment of the upper atmosphere was from solar flare X rays [Peterson and Winckler, 1959; Chubb, Friedman, and Kreplin, 1960] rather than electrons. This required a flare in progress at the time of the burst of radiation. Indeed, a flare of importance 3 was near the end of its development at this time, having started at 2350 UT on June 26, 1960. This flare reached its maximum development at 0010 UT on June 27, 1960, and ended at 0200 UT, some

TABLE 1

Riometer Site	Geographic Latitude	Geographic Longitude	Geomagnetic Latitude	Solar Elevation
Thule, Greenland	76°33' N	68°47' W	86.5°	13°
Resolute Bay, Canada	74°42' N	94°54' W	82.5°	17°
Coral Harbour, Canada	64°07' N	83°14' W	74.5°	8°
Fort Churchill, Canada	58°50' N	94°12' W	68.4°	9°
Trondheim, Norway	68°45' N	16°27' E	66.6°	7°
Longyearbyen, Norway	69°58' N	23°20' E	66.5°	10°
Cape Jones, Canada	54°38' N	79°50' W	65.4°	0°
Umeå, Sweden	67°50' N	20°25' E	65.1°	7°
College, Alaska	64°44' N	147°39' W	64.7°	34°
Wentzville, Alaska	61°58' N	151°11' W	61.5°	37°
St. John's, Canada	48°06' N	77°46' W	59.0°	-5°
St. Lawrence, Canada	45°24' N	75°54' W	56.4°	-8°

minutes after the burst of radiation. A sudden ionospheric disturbance (Hakura, private communication) of importance 2+, apparently related to this flare, was observed to start at 0303 UT on June 27, but it ended about 67 minutes later. The next disturbance of this type to occur began at 0419 UT, too late to account for the ionospheric absorption observed to begin at 0146 UT.

The solar X-ray interpretation is ruled out most conclusively, however, by considering the observations obtained with the high latitude riometers operating at the time of the event. It is seen from Table 1 that at four northern sites (Thule, Resolute Bay, Coral Harbour, and Fort Churchill) the solar elevation angle was at least equal to or greater than the elevation at the northern European riometers where the increased absorption was observed. In view of this, these riometers would have shown increased absorption if the ionospheric disturbance resulted from a burst of solar X rays. The records of these instruments did not, however, show any detectable absorption setting in at the time in question. This negative result rules out the solar X-ray interpretation; in addition, it in-

dicates that the electron bombardment at the sudden commencement does not represent another polar-cap phenomena similar to that found with solar protons [cf., e.g., Reid and Leinbach, 1959].

The geographical distribution of riometer sites where the increased absorption was definitely observed, as well as the fact that it was not observed at high latitudes, suggests that the electrons incident on the atmosphere were confined to a ringlike path around the polar cap in the vicinity of the auroral zone. If this is true, the increased absorption would be expected to be detectable at Cape Jones and Fort Churchill, Canada, respectively south of and in the auroral zone.

There is some slight evidence from the riometer record at Cape Jones that such was the case. The absence of increased absorption at Fort Churchill, with detection conditions comparable with those where the absorption was observed, indicates that the electron bombardment did not extend as far north as the auroral zone. An alternative explanation, and possibly preferable in view of the lack of definite absorption at Cape Jones and at stations farther south such

as Val d'Or and Ottawa, is that the electron bombardment did not occur in a complete ring around the earth, and may even have been limited to several reasonably defined patches.

Discussion. The burst of X rays accompanying the magnetic storm on June 27 began rapidly with the sudden commencement and followed roughly the variations of the horizontal component H for the first 8–10 minutes. No further increases in counting rate were observed during the balloon flight, telemetry having been continuous until 1000 UT, even though strong magnetic activity set in shortly after the sudden commencement. A Forbush decrease of cosmic-ray intensity was evident, however, reaching an amplitude of about 5 per cent in an hour after the sudden commencement before becoming masked by the slow descent of the balloon.

The present event differs from that observed by *Anderson* [1958] on August 29, 1957, in several respects. First, the amplitude of the event was about 3.5 times greater; in addition, the duration of the X-ray burst was only a few minutes instead of several hours after the onset of the magnetic storm. Further, the flare that gave rise to the magnetic storm did not produce a solar proton event such as in August, 1957.

Satellite observations of the reduction in intensity of trapped electrons in the outer radiation zone with magnetic storm activity [*Rothwell and McIlwain*, 1960] have been interpreted as resulting from the dumping of part of the contents of the zone. The manner in which these observations are obtained, from a number of passes of the satellite over areas under the lower portions of the outer zone, makes it difficult to determine details of the dumping process itself, such as the temporal and spatial variations. In this respect, balloon and riometer observations supplement the satellite data.

On the assumption that the X-ray bombardment observed on June 27 represents an example of this process, a rough comparison can be made between the number of electrons impinging on 1 square centimeter of the atmosphere near the auroral zone during the event and the contents of the volume of the outer zone connected to this area. In this connection, the balloon observations indicate that the principal bombardment lasted about 5 minutes and was fairly symmetrical in time. Taking 50 kev as an estimate of the X-ray energy at the detec-

tor, the peak counting rate in the interval 0148–0149 UT corresponds to a flux of $2\text{--}6 \times 10^7$ electrons/cm²-sec with energies greater than 50 kev. This estimate, of course, depends on the electron spectrum assumed incident at the top of the atmosphere; a spectrum of the type for auroral zone X rays [*Anderson*, 1960] was used in calculating this flux on the assumption that the physical processes are not too different. As a result, the X-ray burst was produced by roughly $3\text{--}9 \times 10^9$ electrons with energy greater than 50 kev hitting 1 square centimeter of the atmosphere.

This figure is comparable to the $10^8\text{--}10^{11}$ electrons with energies greater than 20 kev trapped within the volume of the radiation zone having 1 square centimeter base in the auroral zone and located along the lines of force connecting the two hemispheres. In addition, this estimate is roughly the same order of magnitude as the daily flux of 10^{10} electrons/cm² with energies greater than 25 kev which are lost in the atmosphere in the auroral zone [*Anderson*, 1960].

It has been suggested by *Anderson* [1960] that a terrestrial acceleration process must occur regularly, even in the absence of pronounced magnetic activity, to replenish the loss to the outer zone represented by the steady, daily leakage over the auroral zone. On the assumption that the brief, intense electron bombardment at the time of the sudden commencement of a geomagnetic storm involves dumping from the outer zone, it represents another drain, although smaller in magnitude considering the relatively long times between magnetic storms of this type, which must be met by the acceleration process. While the daily loss of electrons to the auroral zone would impose certain requirements on this mechanism, more stringent limitations would be imposed by sudden commencements of the type reported here when the geographical extent of the electron bombardment and the extremes in the time distribution between these events become known more fully.

Another possible interpretation of the appearance of X rays at the time of a sudden commencement geomagnetic storm is that the electrons responsible for the X rays were released from a solar plasma cloud on its interaction with the geomagnetic field. In view of the localization of the electron bombardment near the auroral zone, a means of channeling the

icles into this region would be required. particle precipitation near the auroral zone, either a complete or patchy ring, would be related to the structure of the plasma clouds. In addition, the frequency of events such as this would depend only on solar activity and not be connected with acceleration processes taking place in the vicinity of the earth.

Acknowledgments. The balloon observations were supported in part by the joint program of the U. S. Office of Naval Research and the U. S. Atomic Energy Commission. The riometer observations at College, Alaska, and Thule, Greenland, were supported in part by the Electronic Research Directorate, Air Force Command and Control Development Division, and by the National Science Foundation. The riometer observations at Kiruna, Sweden, were supported in part by the European Office of Air Research and Development Command under contract AF 61(514)-34; those at Alta and Harstad, Norway, were supported in part under contract AF 61(052)-08. One of us (R. R. B.) is indebted to C. D. Anger, S. Evans, and W. B. Hughes for their able assistance in carrying out the balloon observations. In addition, we are indebted to the High Altitude Observatory and the U. S. Coast and Geodetic Survey for solar and magnetic data. Dr. C. Gordon Cole, Central Radio Propagation Laboratory, U. S. National Bureau of Standards, kindly provided riometer data from Kenai and Skwentna, Alaska. Dr. Y. Hakura of the Hiraiso Radio Observatory provided information on SID effects. Mr. John Allen Jones and Mr. Jamie Chapman of the GRD Geopole Station, Thule, Greenland, were indebted for the operation of the

Thule riometer. We thank Dr. B. Hultqvist for reading and discussing the manuscript.

REFERENCES

- Anderson, K. A., Soft radiation events at high altitude during the magnetic storm of August 29-30, 1957, *Phys. Rev.*, **111**, 1397, 1958.
- Anderson, K. A., Balloon observations of X rays in the auroral zone II, *J. Geophys. Research*, **65**, 3521-3538, 1960.
- Chubb, T. A., H. Friedman, and R. W. Kreplin, Measurements made of high-energy X rays accompanying three class 2+ solar flares, *J. Geophys. Research*, **65**, 1831-1832, 1960.
- Little, C. G., and H. Leinbach, The riometer—a device for the continuous measurement of ionospheric absorption, *Proc. IRE*, **47**, 315, 1959.
- Peterson, L. E., and J. R. Winckler, Gamma-ray burst from a solar flare, *J. Geophys. Research*, **64**, 697-707, 1959.
- Reid, G. C., and H. Leinbach, Low-energy cosmic ray events associated with solar flares, *J. Geophys. Research*, **64**, 1801-1805, 1959.
- Rothwell, P., and C. E. McIlwain, Magnetic storms and the Van Allen radiation belts: observations from Satellite 1958e (Explorer IV), *J. Geophys. Research*, **65**, 799-806, 1960.
- Van Allen, J. A., Direct detection of auroral radiation with rocket equipment, *Proc. Natl. Acad. U. S.*, **43**, 67, 1957.
- Winckler, J. R., L. Peterson, R. Arnoldy, and R. Hoffman, X-rays from visible aurorae at Minneapolis, *Phys. Rev.*, **110**, 1221, 1958.
- Winckler, J. R., Balloon study of high altitude radiations during the International Geophysical Year, *J. Geophys. Research*, **65**, 1331-1359, 1960.

(Manuscript received January 7, 1961.)

The Steady State of the Chapman-Ferraro Problem in Two Dimensions

J. W. DUNGEY

*Ionosphere Research Laboratory, The Pennsylvania State University
University Park, Pennsylvania*

Abstract. The steady state of the Chapman-Ferraro problem is formulated in mathematical terms, but the three-dimensional problem presents little hope of an analytical solution, and it is not even obvious how to compute the solution. The two-dimensional problem is reduced to a standard potential problem and solved exactly. The solution is used to obtain an indication of the error in the flat-faced approximate model that has been used previously in three dimensions.

Formulation of the problem. From consideration of the solar origin of magnetic storms Chapman and Ferraro were led to make a theoretical study of what happens when a stream of neutral ionized gas approaches a magnetic dipole. They were mainly concerned with the early phase when the front of the stream is approaching and wrapping around the dipole. The problem in which the stream continues to flow for a very long time is also of interest, because, even in the absence of a solar wind, the orbital motion of the earth sets up the same situation [Dungey, 1958]. The earth's magnetic field is screened into a cavity (see Fig. 1) into which the streaming plasma cannot penetrate, the surface layer being thin (a few kilometers), but the steady state is unstable to the formation of traveling surface waves. Parker [1958] has pointed out that the double streaming just outside the cavity is likely to be another source of instability. In spite of these instabilities, it is interesting to try and work out the steady-flow problem, assuming a uniform stream. It is interesting to know where the lines of force from the neutral points N on the surface (Chapman and Ferraro's horns) strike the earth, and the problem is of some mathematical interest as a free-boundary problem, which cannot be solved by any known technique. Since the work was done, however, an interplanetary magnetic field of a few gammas under quiet conditions has been discovered by Pioneer [Coleman, Davis, and Sonett, 1960] and this seems to put an end to its applicability to the earth.

By neglecting all but the essential effects and taking the surface of the cavity to be perfectly

thin, the problem can be formulated so that the boundary conditions are very simple [Dungey, pp. 144-145, 1958]. The field is supposed to be completely screened, so that as one condition the field at the boundary should be tangential to the boundary. It is a good approximation that particles from the stream are specularly reflected at the surface. It is also permissible to neglect the thermal velocities of the particles in the stream, since the mass is concentrated in the ions and their thermal velocities are small compared to the stream speed. Thermal motion is important in closing the cavity on the leeward side, and, when it is neglected, the cavity extends to infinity behind the earth. Further, we neglect the interaction between incoming particles and those which have bounced off the cavity surfaces; the mean free path is very large compared with the cavity, but the double stream instabilities may cause a strong interaction as suggested by Parker.

With these approximations the pressure of the stream on the surface can be calculated from the impulse on the particles and is $nmv^2 \cos^2 \chi$, where n is the number density of ions, m their mass, v the speed of the stream, and χ the angle between the velocity of the stream and the inward normal of the surface (χ must be acute all over the surface). This pressure must be balanced by the pressure of the magnetic field on the inside of the surface, which is $H^2/8\pi$. The boundary condition is therefore

$$H = (8\pi nm)^{1/2} v \cos \chi \quad (1)$$

in which n , m , and v are constants. The problem is to find the field inside the cavity which con-

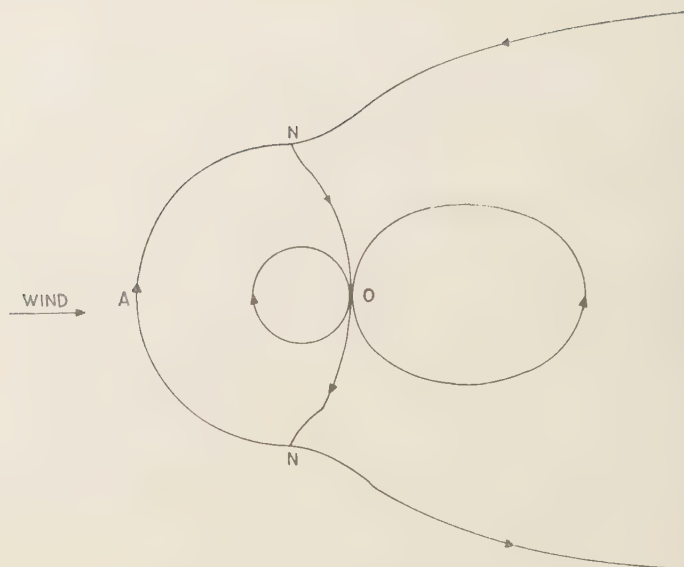


Fig. 1. Two-dimensional representation of the steady-state Chapman-Ferraro problem.

tains no source of magnetic field except for one dipole and which satisfies the above boundary conditions on some unknown surface. The constants of the problem can be scaled so that the only parameter to be fixed is the angle between the velocity of the stream and the dipole. If the other constants of the stream are changed, the cavity changes in size, but not in shape, so that the field at equivalent points on the surfaces change in proportion to $n^{1/2} u$; then, if the dipole moment is kept the same, any length must change in proportion to $n^{1/6} u^{1/3}$.

The simplest choice for the direction of the stream is perpendicular to the dipole (this formulation fails if it is parallel to the dipole, it is an event unlikely in the application to the earth). There are then two planes of symmetry, but this does not make a great simplification—it does not reduce the number of dimensions.

There appears to be little prospect of obtaining an analytical solution of this three-dimensional problem and it is not even clear how to compute it, since the surface wave instability would appear in a time-dependent solution. The equivalent two-dimensional problem, however, with a line dipole, is amenable to the method of complex potentials and it turns out that the free boundary becomes a very simple known boundary in

potential space. The solution of the two-dimensional problem is presented in the next section and the position of the horns is briefly discussed in the final section.

Solution of the two-dimensional problem. Let the stream flow in the x direction and let the dipole point in the y direction. Denoting the surface of the cavity by S , the angle χ between the inward normal and the x direction is given by

$$\cot \chi = (dy/dx)_S \quad (2)$$

Because of the way in which the problem scales, we may put $8\pi nmu^2 = 1$, which is equivalent to taking $H = 1$ at the apex A in Figure 1. With this choice, and using (2), the boundary condition (1) becomes

$$H = [1 + (dx/dy)_S^2]^{-1/2} \quad (3)$$

Now, when a scalar potential is introduced such that $H = \nabla V$, (3) can be integrated along S . Referring to Figure 1 for the direction of the field the most convenient form is

$$\left. \begin{aligned} V &= -y - 2 & -2 < y < -1 \\ V &= y & -1 < y < 1 \\ V &= 2 - y & 1 < y < 2 \end{aligned} \right\} \text{ on } S \quad (4)$$

It may be noted that the same integration may be performed in the three-dimensional case with stream velocity perpendicular to the dipole, not only in the plane of symmetry containing the dipole (the plane of Fig. 1). In both instances it is seen that the cavity lies between the planes $y = \pm 2$ and touches them as $x \rightarrow \infty$.

The complex potential also involves the stream function or vector potential U , which is constant on a line of force. Since the field on the boundary is tangential, U is constant over the whole boundary, and, for convenience, its value on the boundary will be taken as zero. Since this gives $U = 0$ at N , the line of force from N to the dipole also has $U = 0$, and it will be seen that this line of force is determined by the analysis in a simple way.

The complex potential is written

$$\Phi = U + iV$$

and we put

$$Z = x + iy$$

The principle of the method is that for Φ to satisfy Laplace's equation in x and y , Φ must be an analytic function of Z . If the boundary were fixed (known in x and y), we would look for a $\Phi(Z)$ which satisfied the boundary conditions there. In the present problem, however, the boundary is not known in ordinary space, but is known in U and V . Now, if Φ is an analytical function of Z , Z must be an analytical function of Φ , so this problem is solved by finding a function $Z(\Phi)$ to fit the boundary conditions. It must be remembered that there is one source of field inside the cavity, the line-dipole. At the dipole the potentials have singularities such that, if the dipole is at Z_0 , Φ behaves like $(Z - Z_0)^{-1}$. The dipole, therefore, corresponds to the behavior of $Z(\Phi)$ as $\Phi \rightarrow \infty$; Z must behave like $M\Phi^{-1}$, where M is the dipole moment. Thus in the equivalent potential problem in which Φ and Z are interchanged, there must be a dipole moment but no net charge.

The boundary in Φ -space is the strip $U = 0$, $-1 < V < 1$ and on it y is known from (4); x is determined by the solution. Because the boundary contains neutral points, it is singular, and Φ is a double-valued function of V on the boundary. The points, $U = 0$, $V = \pm 1$, which correspond to the neutral points in ordinary space, are branch points of $Z(\Phi)$. The inside of the

cavity maps one to one onto the whole of the Φ -space, and $Z(\Phi)$ is continuous everywhere except on the boundary, where (4) gives two values of y for each point. It is convenient to take $U > 0$ in the part of the cavity bounded by the lines of force AN and NO . Then from Figure 1, $U < 0$ in the remainder of the cavity which extends to infinity, so that these two parts of the cavity correspond to the half planes in Φ -space $U \gtrless 0$. Now since the boundary condition on AN is $V = y$, the boundary value $y = V$ belongs to that side of the strip which faces the half-plane $U > 0$. The other side has $y = -V - 2$ when $-1 < V < 0$, and $y = 2 - V$ when $0 < V < 1$, the discontinuity at $V = 0$ corresponding to $x \rightarrow \infty$.

The problem now has the form of a potential problem thinking of $y(\Phi)$ as the potential. It is convenient to treat it as the sum of two parts: (a) has the boundary values on $U = 0$, $-1 < V < 1$

$$y = 1 \quad \text{when} \quad 0 < V < 1$$

$$y = -1 \quad \text{when} \quad -1 < V < 0$$

for both sides of the strip. This contains the dipole source. (b) has the boundary values on $U = 0$, $-1 < V < 1$

$$y = \alpha(V - 1) \quad \text{when} \quad 0 < V < 1$$

$$y = \alpha(V + 1) \quad \text{when} \quad -1 < V < 0$$

where $\alpha = +1$ for the side facing $U > 0$ and $\alpha = -1$ for the side facing $U < 0$. This has a quadrupole nature (thinking of y as the potential).

The solution for (a) is

$$Z = -2\pi^{-1}sh^{-1}\Phi^{-1} \quad (5)$$

To see this, and to obtain x and y , we expand it to

$$sh(\pi x/2) \cos(\pi y/2) = -U(U^2 + V^2)^{-1} \quad (6)$$

$$ch(\pi x/2) \sin(\pi y/2) = V(U^2 + V^2)^{-1} \quad (7)$$

When $U = 0$, (6) is satisfied by $y = \pm 1$ and (7) shows that the sign is the same as that of V in agreement with conditions (a). Then from (7)

$$x = \pm 2\pi^{-1}ch^{-1}|V|^{-1} \quad (8)$$

the sign being determined as follows. In the potential problem corresponding to conditions

(a) the extreme values (in fact, the only values) of y on the boundary are ± 1 . Consequently, in this problem, $-1 < y < 1$ anywhere in Φ -space and hence $\cos(\pi y/2) > 0$. Then from (6) it is seen that x has the opposite value to U , and for continuity this must be true also on the boundary.

The part of $U = 0$ with $|V| > 1$ corresponds to the line of force NO , which we wish to investigate. Here (8) does not give a real value for x , and (6) and (7) can only be satisfied by $x = 0$. Then $y = 2\pi^{-1} \sin^{-1} V^{-1}$. The dipole is determined by (5) with Φ large giving $\Phi \approx -2\pi^{-1} Z^{-1}$.

The part (b) is given by

$$Z = \pi^{-1}[(1 + i\Phi) \log(1 - i\Phi^{-1}) + (1 - i\Phi) \log(1 + i\Phi^{-1}) - 2] \quad (9)$$

When Φ is large, this is $O(\Phi^{-2})$, which shows its quadrupole nature.

On $U = 0$ the value of y depends on the imaginary parts of the logarithms. The imaginary part of $\log(1 - V^{-1})$ is zero except when $0 < V < 1$. Now $\log(1 - i\Phi^{-1})$ has branch points at $\Phi = 0$ and $\Phi = i$. When Φ goes anticlockwise around the branch point at $\Phi = 0$, $\log \Phi$ changes by $2\pi i$ and hence $\log(1 - i\Phi^{-1})$ changes by $-2\pi i$, so that we must take it as $-\pi i$ on the side facing $U > 0$ and πi on the other side. Similarly the imaginary part of $\log(1 + V^{-1})$ vanishes except when $-1 < V < 0$. Here it must be taken as πi on the side facing $U > 0$ and $-\pi i$ on the other side. Thus conditions (b) are satisfied.

The values of x on $U = 0$ obtained from (9) are

$$x = \pi^{-1}[(1 - V) \log(V^{-1} - 1) + (1 + V) \log(V^{-1} + 1) - 2] \quad \text{when } 0 < V < 1$$

$$x = \pi^{-1}[(1 - V) \log(1 - V^{-1}) + (1 + V) \log(-1 - V^{-1}) - 2] \quad \text{when } -1 < V < 0$$

$$x = \pi^{-1}[(1 - V) \log(1 - V^{-1}) + (1 + V) \log(1 + V^{-1}) - 2] \quad \text{when } V > 1 \text{ or } V < -1$$

When $V > 1$ or $V < -1$, (9) gives $y = 0$.

Combining (a) and (b), we get for that part of the surface with y positive $0 < V < 1$,

$$y = V \quad \text{with} \\ \pi x = (1 - V) \log(V^{-1} - 1) + (1 + V) \log(V^{-1} + 1) - 2ch^{-1}V^{-1} - 2 \quad (10)$$

and

$$y = 2 - V \quad \text{with} \\ \pi x = (1 - V) \log(V^{-1} - 1) + (1 + V) \log(V^{-1} + 1) + 2ch^{-1}V^{-1} - 2 \quad (11)$$

For the line of force NO

$$V > 1, \pi y = 2 \sin^{-1} V^{-1} \quad \text{with} \\ \pi x = (1 - V) \log(1 - V^{-1}) + (1 + V) \log(1 + V^{-1}) - 2 \quad (12)$$

These curves are plotted in Figure 2 and are found to have the general shape expected from Figure 1.

The lines of force NO. It is interesting to know where the line of force NO hits the earth in the three-dimensional problem (before the discovery of an interplanetary field this might have been interpreted as the auroral zone). Previous estimates have been obtained from Chapman and Ferraro's model with a flat-faced stream, the solution being obtained by introducing an image dipole. This gives a geomagnetic latitude of about 77° [Dungey, p. 147, 1958]. Now, though the two-dimensional solution can not give a value for this latitude directly, it can give an indication of the error in the flat-faced model. The two-dimensional flat-faced model is very simple, and will be given now and compared later with the exact two-dimensional solution.

The vector potential for two unit dipoles at the origin and $(-2a, 0)$ is

$$A = \frac{x}{x^2 + y^2} + \frac{x + 2a}{(x + 2a)^2 + y^2} \quad (13)$$

The lines of force corresponding to those shown in Figure 2 are given by $A = 0$ and are the flat-faced $x = -a$ and the circle on the diameter



Fig. 2. Curves obtained from equations 10, 11, and 12.

ining the dipoles, $x^2 + y^2 + 2ax = 0$. Near the origin, therefore, this line of force is characterized by $y^2/x \rightarrow -2a$. If an appropriate value for a can be found, this can be compared with the exact solution. In the exact solution for ON far 0, V is large and (12) can be approximated

$$\begin{aligned} & \approx 2V^{-1} \\ & \pi = (V-1)(V^{-1} + \frac{1}{2}V^{-2} + \dots) \\ & \quad + (V+1)(V^{-1} - \frac{1}{2}V^{-2} + \dots) \\ & -2 \approx -\frac{1}{3}V^{-2} \end{aligned}$$

ing

$$y^2/x \rightarrow -12/\pi$$

One plausible method of comparison would be to take a to be the same as OA in the exact solution, that is, $a = 1.08$. Then the flat-faced model gives a value of y^2/x not much greater than half its true value, indicating that the latitude obtained is considerably too low.

An alternative method of comparison is to

match the field strengths at A and $(0, -a)$ remembering that the dipole in the exact solution has strength $2/\pi$ from (5). The field strength at $(0, -a)$ from (13) is $2a^{-2}$, and so we must put $2a^{-2} = \pi/2$ giving $-a = 2/\sqrt{\pi}$ and the limiting value of y^2/x is too small by the factor $\sqrt{\pi}/3 \approx 0.6$. These two methods of comparison agree well, and it can be concluded that the latitude obtained from the flat-faced model is too small, so that the true latitude must be substantially greater than 77° .

Acknowledgment. The research reported in this paper has been sponsored by the National Science Foundation under Grant 5939 and, in part, by the Geophysics Research Directorate of the Air Force Cambridge Research Center under contract AF19(604)-3875.

REFERENCES

- Coleman, P. J., Leverett Davis, and C. P. Sonett, *Phys. Rev. Letters*, **5**, 43, 1960.
- Dungey, J. W., *Cosmic Electrodynamics*, Cambridge University Press, 1958.
- Parker, E. N., *Physics of Fluids*, **1**, 171, 1958.

(Manuscript received January 31, 1961.)

Radiation from a Current Filament above a Homogeneous Earth, with Application to Micropulsations

P. F. LAW AND B. M. FANNIN

*Electrical Engineering Research Laboratory, The University of Texas
Austin, Texas*

Abstract. If electromagnetic radiations from ionospheric current systems are considered sources of micropulsations, it becomes instructive to examine the following model problem: What are the electric and magnetic fields in the vicinity of a plane, homogeneous earth due to a line current source above this model earth? As the micropulsation frequencies are quite low (order of 1 cps), the nearfield considerations become significant. The solution for this near field problem is presented and the fields at the earth's surface are evaluated for a set of parameter values chosen to approximate an ionospheric source at a typical micropulsation frequency.

Introduction. There is a form of energy incident upon the earth's surface that manifests itself in the form of observable time and space variations of the electric and magnetic fields. Certain of these variations are termed 'micropulsations.' The term 'geomagnetic micropulsations' is meant to refer to those regular, naturally occurring, fluctuations of the earth's magnetic field having amplitudes of the order of a gamma 10^{-6} gauss or 10^{-9} weber/square meter) and periods from about 0.1 second to 10 minutes. The associated fluctuations of the electric field may be termed 'micropulsations' in the earth's electric field.

Benkova [1953], Hasegawa [1960], Jacobs and Linno [1960], and Tamm [1944] have indicated that the transient behavior of the geomagnetic field may be related to electric currents in the ionosphere or outer atmosphere. The assumption of an ionospheric source leads at once to the conclusion that the earth is in the near field of such a radiator. It is true, in fact, that, at micropulsation frequencies, the ionosphere-earth distance is only a small fraction (0.01 or less) of wavelength.

The assumption that micropulsations may be represented locally as plane waves incident on the earth leads to the conclusion that the downward wave impedance in a homogeneous, conducting earth is

$$\sqrt{\frac{\omega \mu_e}{\sigma_e}} / 45^\circ$$

where ω , μ_e , and σ_e are the radian frequency, earth permeability, and earth conductivity, respectively. Further, such a hypothesis predicts a very small vertical component of magnetic field variation. This latter conclusion has led some authors to theorize that the observed vertical component (usually of appreciable size when compared with the horizontal) is due solely to inhomogeneities in the earth's structure or slanted stratification planes within the earth. Similar arguments are used to explain the observed variations in wave impedance at a fixed frequency. The results of this study indicate that near field considerations give rise to an appreciable vertical component of magnetic field variation. The downward wave impedance of the homogeneous earth remains, however, substantially that value predicted by the plane wave approach.

A line source parallel to a plane, homogeneous earth. In this paper the authors consider the problem of finding the fields associated with a filamentary conductor of infinite length, carrying an in-phase current above a plane, homogeneous earth. A right-hand rectangular coordinate system is established so that the x - z plane is the plane of the earth and the positive y direction is upward. The line source is taken to be parallel to the z axis, and the current flow is in the negative z direction. The current will be assumed to vary harmonically and be taken to have the time functional form $e^{i\omega t}$ of phasor notation. The time variations are suppressed in the expres-

functional form. Further, it will be assumed that the magnetic field is parallel to the x - y plane and that the electric field is parallel to the z axis.

The field at $(x, 0 < y < h, z)$ is due to a summation of the source effect (primary field) and the image effect (secondary field). These may be written (see *Law and Fannin* [1960]) as

Primary field components ($0 < y < h$):

$$\begin{aligned} \bar{E}_1 &= \frac{\eta I}{2\lambda} \int_{-\infty}^{\infty} (0, 0, 1) \\ &\cdot e^{-jk[Sx-C(y-h)]} \frac{dS}{C} \quad (1) \end{aligned}$$

$$\begin{aligned} \bar{H}_1 &= \frac{I}{2\lambda} \int_{-\infty}^{\infty} (-C, -S, 0) \\ &\cdot e^{-jk[Sx-C(y-h)]} \frac{dS}{C} \quad (2) \end{aligned}$$

Secondary field components ($0 < y < h$):

$$\begin{aligned} \bar{E}_2 &= \frac{\eta I}{2\lambda} \int_{-\infty}^{\infty} (0, 0, 1) \rho(S) \\ &\cdot e^{-jk[Sx+C(y+h)]} \frac{dS}{C} \quad (3) \end{aligned}$$

$$\begin{aligned} \bar{H}_2 &= \frac{I}{2\lambda} \int_{-\infty}^{\infty} (C, -S, 0) \rho(S) \\ &\cdot e^{-jk[Sx+C(y+h)]} \frac{dS}{C} \quad (4) \end{aligned}$$

the function $\rho(S)$ that appears in the secondary field components is the Fresnel reflection coefficient of the earth. This coefficient for horizontally polarized incidence is given as

$$\begin{aligned} \rho(S) &= \frac{C - \sqrt{S_0^2 - S^2}}{C + \sqrt{S_0^2 - S^2}} = -1 \\ &+ \frac{2C}{C + \sqrt{S_0^2 - S^2}} \quad (5) \end{aligned}$$

where $C = \cos \theta$, $S = \sin \theta$, h = radiator height (see fig. 1),

$$\begin{aligned} S_0^2 &= \epsilon^*/\epsilon, \quad \epsilon^* = \epsilon_s + \sigma_s/j\omega, \\ \eta &= \sqrt{\mu_0/\epsilon_0}, \quad k = \omega \sqrt{\mu_0\epsilon_0}, \end{aligned}$$

and $\lambda = 2\pi/k$.

The total electric and magnetic fields result from an addition of the primary and secondary effects:

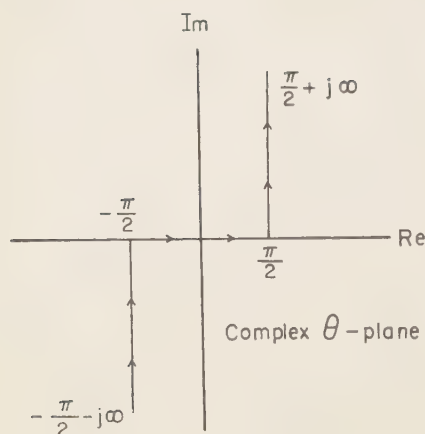
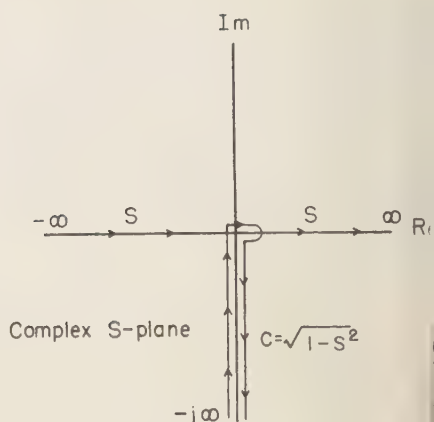
$$\begin{aligned} \bar{E}(x, 0 < y < h) &= \frac{\eta I}{2\lambda} \int_{-\infty}^{\infty} (0, 0, 1) \\ &\cdot e^{-jk[Sx-C(y-h)]} \frac{dS}{C} \\ &- \frac{\eta I}{2\lambda} \int_{-\infty}^{\infty} (0, 0, 1) e^{-jk[Sx+C(y+h)]} \frac{dS}{C} \\ &+ \frac{\eta I}{2\lambda} \int_{-\infty}^{\infty} (0, 0, 1) \frac{2C}{C + \sqrt{S_0^2 - S^2}} \\ &\cdot e^{-jk[Sx+C(y+h)]} \frac{dS}{C} \quad (6) \end{aligned}$$

$$\begin{aligned} \bar{H}(x, 0 < y < h) &= \frac{I}{2\lambda} \int_{-\infty}^{\infty} (-C, -S, 0) \\ &\cdot e^{-jk[Sx-C(y-h)]} \frac{dS}{C} \\ &- \frac{I}{2\lambda} \int_{-\infty}^{\infty} (C, -S, 0) e^{-jk[Sx+C(y+h)]} \frac{dS}{C} \\ &+ \frac{I}{2\lambda} \int_{-\infty}^{\infty} (C, -S, 0) \frac{2C}{C + \sqrt{S_0^2 - S^2}} \\ &\cdot e^{-jk[Sx+C(y+h)]} \frac{dS}{C} \quad (7) \end{aligned}$$

The first two integrals may be evaluated, the results being given in terms of Hankel functions as follows:

$$\begin{aligned} \bar{E}(x, 0 < y < h) &= \bar{a}_x \frac{\eta I \pi}{2\lambda} H_0^{(2)}(k\rho_1) \\ &- \bar{a}_x \frac{\eta I \pi}{2\lambda} H_0^{(2)}(k\rho_2) \\ &+ \frac{\eta I}{2\lambda} \int_{-\infty}^{\infty} (0, 0, 1) \frac{2C}{C + \sqrt{S_0^2 - S^2}} \\ &\cdot e^{-jk[Sx+C(y+h)]} \frac{dS}{C} \quad (8) \end{aligned}$$

$$\begin{aligned} \bar{H}(x, 0 < y < h) &= \bar{a}_{\phi 1} \frac{jI \pi}{2\lambda} H_1^{(2)}(k\rho_1) \\ &- \bar{a}_{\phi 2} \frac{jI \pi}{2\lambda} H_1^{(2)}(k\rho_2) \\ &+ \frac{I}{2\lambda} \int_{-\infty}^{\infty} (C, -S, 0) \frac{2C}{C + \sqrt{S_0^2 - S^2}} \\ &\cdot e^{-jk[Sx+C(y+h)]} \frac{dS}{C} \quad (9) \end{aligned}$$

(a) PATH OF θ (b) PATHS OF S AND C Fig. 2. Variations of θ , S , and C .

where (ρ_1, ϕ_1, z) are the cylindrical coordinates of the point of observation relative to the source, and (ρ_2, ϕ_2, z) are the cylindrical coordinates of the point of observation relative to the image.

Note that the terms in equations 8 and 9 containing the Hankel functions represent uniform cylindrical waves emanating from the source and its image. For the complete solution there remains the evaluation of an integral of the form

$$I(x, y) = \int_{-\infty}^{\infty} \frac{2C}{C + \sqrt{S_0^2 - S^2}} \cdot e^{-jk[Sx + C(y+h)]} \frac{dS}{C} \quad (10)$$

In the evaluation of this integral it is necessary to consider how C varies with S . Figure 2 shows a suitable choice of θ and paths of C and S . Figure 2(b) shows that, over most of the range of integration $(-\infty < S < \infty)$,

$$C = \sqrt{1 - S^2} \cong \begin{cases} jS, & S < 0 \\ -jS, & S > 0 \end{cases} \quad (11)$$

Incorporation of these approximations into the integral infers that contributions for small S are relatively insignificant in evaluation of the integral. For the very low frequencies, it may be shown that this is true (see Appendix). Sub-

stitution of these approximations may be made as follows:

$$\begin{aligned} I(x, y) &\cong \int_{-\infty}^0 \frac{2jS}{jS + \sqrt{S_0^2 - S^2}} \cdot e^{-jk[Sx + jS(y+h)]} \frac{dS}{jS} \\ &\quad + \int_0^{\infty} \frac{-2jS}{-jS + \sqrt{S_0^2 - S^2}} \cdot e^{-jk[Sx - jS(y+h)]} \frac{dS}{-jS} \\ &= \frac{2}{S_0^2} \int_0^{\infty} (\sqrt{S_0^2 - S^2} + jS) \cdot e^{-kS(y+h)} [e^{jkSx} + e^{-jkSx}] dS \\ &= \frac{4j}{S_0^2} \int_0^{\infty} S e^{-kS(y+h)} \cos kSx dS \\ &\quad + \frac{2}{S_0^2} \int_0^{\infty} \sqrt{S_0^2 - S^2} \{e^{-k[(y+h) - ix]} + e^{-k[(y+h) + ix]}\} dS \end{aligned} \quad (12)$$

Referring to Erdelyi, Magnus, Oberhettinger, and Tricomi [1954, p. 157] and Erdelyi, Magnus, Oberhettinger, and Tricomi [1953, p. 38], respectively, the two integrals of the expression above may be evaluated. The results may be written as

$$(x, y) \cong \frac{4j}{S_0^2} \left[\frac{(y+h)^2 - x^2}{k^2 \rho_2^4} \right] + j \left\{ \pi \frac{H_1(w) - N_1(w)}{w} + \pi \frac{H_1(\mathbf{w}) - N_1(\mathbf{w})}{\mathbf{w}} \right\} \quad (13)$$

here

$$w = k[x + j(y+h)]S_0.$$

$$\mathbf{w} = k[-x + j(y+h)]S_0.$$

H_1 is a Struve function of order 1.

N_1 is a Bessel function of the second kind of order 1.

With these results, the electric field for the region $0 < y < h$ is readily written as

$$(x, 0 < y < h) \cong \bar{a}_s \frac{\eta I \pi}{2\lambda} H_0^{(2)}(k\rho_1) - \bar{a}_s \frac{\eta I \pi}{2\lambda} H_0^{(2)}(k\rho_2) + \bar{a}_s \frac{\eta I}{2\lambda} \left(\frac{4j}{S_0^2} \right) \left[\frac{(y+h)^2 - x^2}{k^2 \rho_2^4} \right] + \bar{a}_s \frac{j\eta I}{2\lambda} \left\{ \pi \frac{H_1(w) - N_1(w)}{w} + \pi \frac{H_1(\mathbf{w}) - N_1(\mathbf{w})}{\mathbf{w}} \right\} \quad (14)$$

The magnetic field components may then be determined by making use of Maxwell's curl equation $\nabla \times \bar{E} = -\dot{\bar{B}}$.

Evaluation of the fields at the earth's surface. For evaluation of the fields at the earth's surface, certain parameter choices have been made:

$h = 2 \times 10^5$ meters (to approximate the ionospheric current height).

$\sigma_e = 1/8000$ mho/meter (as earth's conductivity).

$\omega = 0.3$ radian/second (as micropulsation frequency).

These choices imply values

$$\begin{aligned} \omega &= \omega \sqrt{\mu_0 \epsilon_0} = 10^{-9} \text{ radian/meter} \\ &= \sqrt{\frac{\epsilon^*}{\epsilon_0}} = \sqrt{\frac{\epsilon_e + \sigma_e / j\omega}{\epsilon_0}} \\ &\cong \sqrt{\frac{\sigma_e}{\omega \epsilon_0}} / -45^\circ = 6860 / -45^\circ \end{aligned}$$

The most expeditious method of numerical evaluation of the various Hankel, Bessel, and Struve functions involved is by digital computer operation on the series representations of these functions.

Figures 3, 4, and 5 show the magnitudes and phases of \bar{E}_z , \bar{H}_z , and \bar{H}_y , respectively, plotted as functions of x . Here x is the distance along the earth's surface from the geographical position of the radiator (see Fig. 1). The magnitudes of the field strengths have been normalized by setting $I/2\lambda = 1$ ampere/meter. The phases are given relative to their respective phases at the source.

Figure 6 shows the magnitude and angle of η_a , the apparent impedance of the earth. This downward wave impedance is defined as the ratio of the tangential components of \bar{E} and \bar{H} at the earth's surface ($\eta_a = \bar{E}_z / \bar{H}_z$).

Superimposed on the plots of true intensities given in Figures 2, 4, and 5 are the values (magnitudes and phases) obtained by considering the incident radiation to have plane phase fronts. At each point, x , along the surface, the propagation direction is taken to be that direction defined by the vector pointing from the source at $y = h$ to the point. The intensities of the incident radiations are taken to be those values that would exist in the absence of the earth. The reflection coefficient for horizontally polarized incidence is given by equation 5. On the several plots, these plane wave results have been modified by scale factors as indicated thereon.

Conclusions. The results of this study indicate that the vertical magnetic component of the micropulsation field at the earth's surface may arise as a result of the near field effects if ionospheric currents are regarded as the radiating sources. The simplified model analysis presented herein predicts a vertical component of magnetic field variation that ranges from 0 to 50 per cent of the horizontal intensity. It seems logical to presume that, for a more extensive current system and a more exact earth model, a vertical component of magnetic field intensity should be present where this component is in part due to near field considerations.

It appears that the apparent impedance of the earth is not a very critical function of distance from the radiating source. Further, it is seen that, for the homogeneous earth model, the apparent impedance is essentially the value that

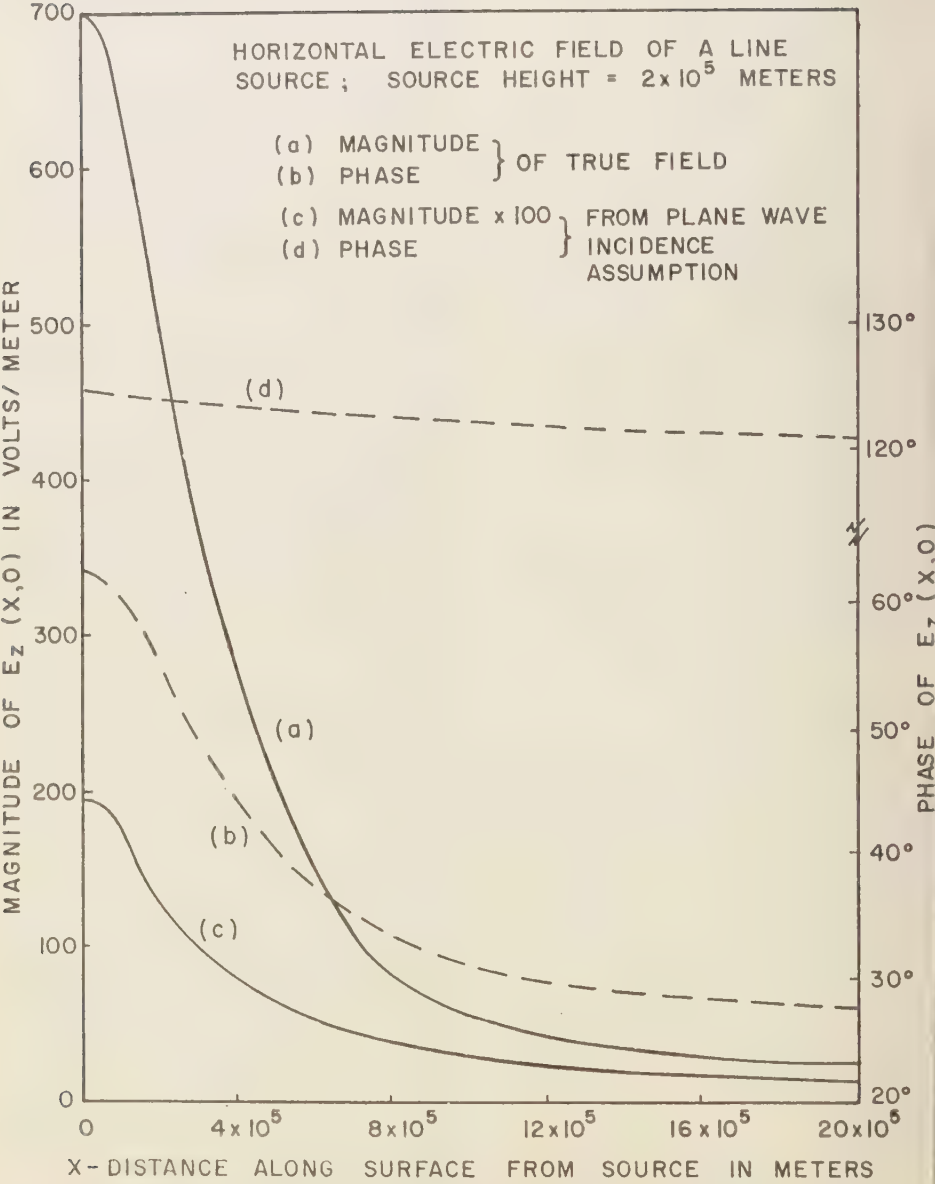


Fig. 3. Magnitude and phase of horizontal electric field intensity, $\vec{E}_z(x, 0)$.

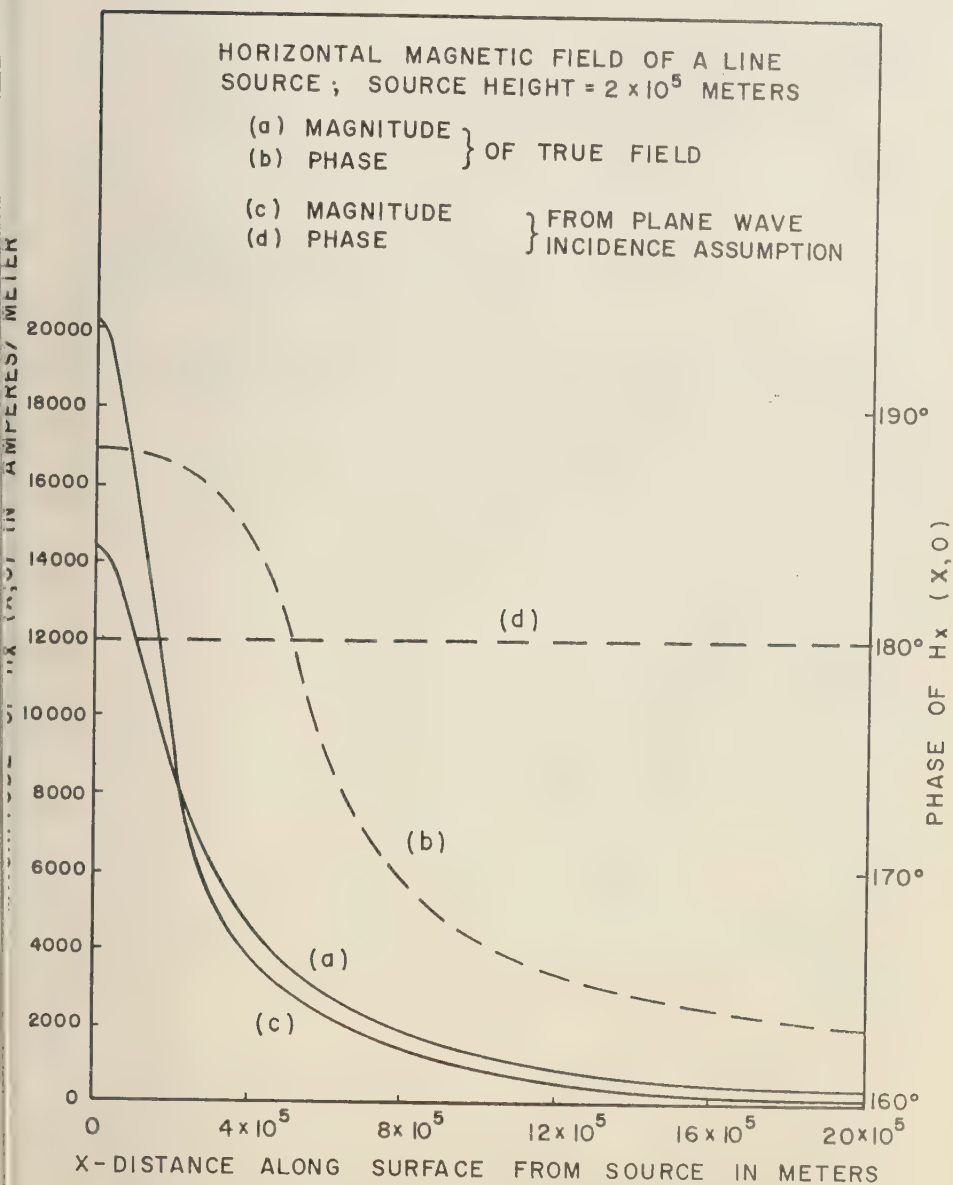


Fig. 4. Magnitude and phase of horizontal magnetic field intensity, $\vec{H}_x(x, 0)$.

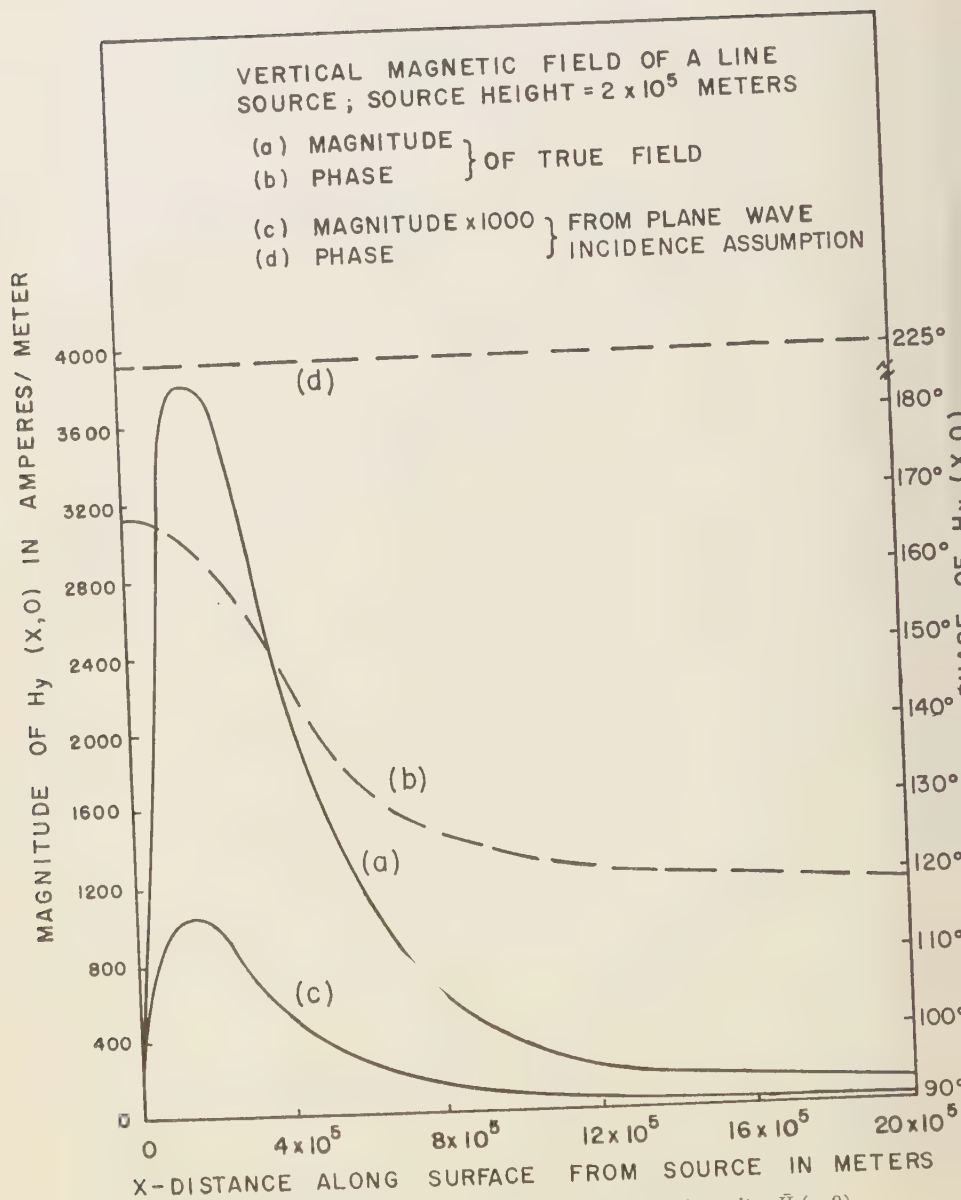


Fig. 5. Magnitude and phase of vertical magnetic field intensity, $\bar{H}_y(x, 0)$.

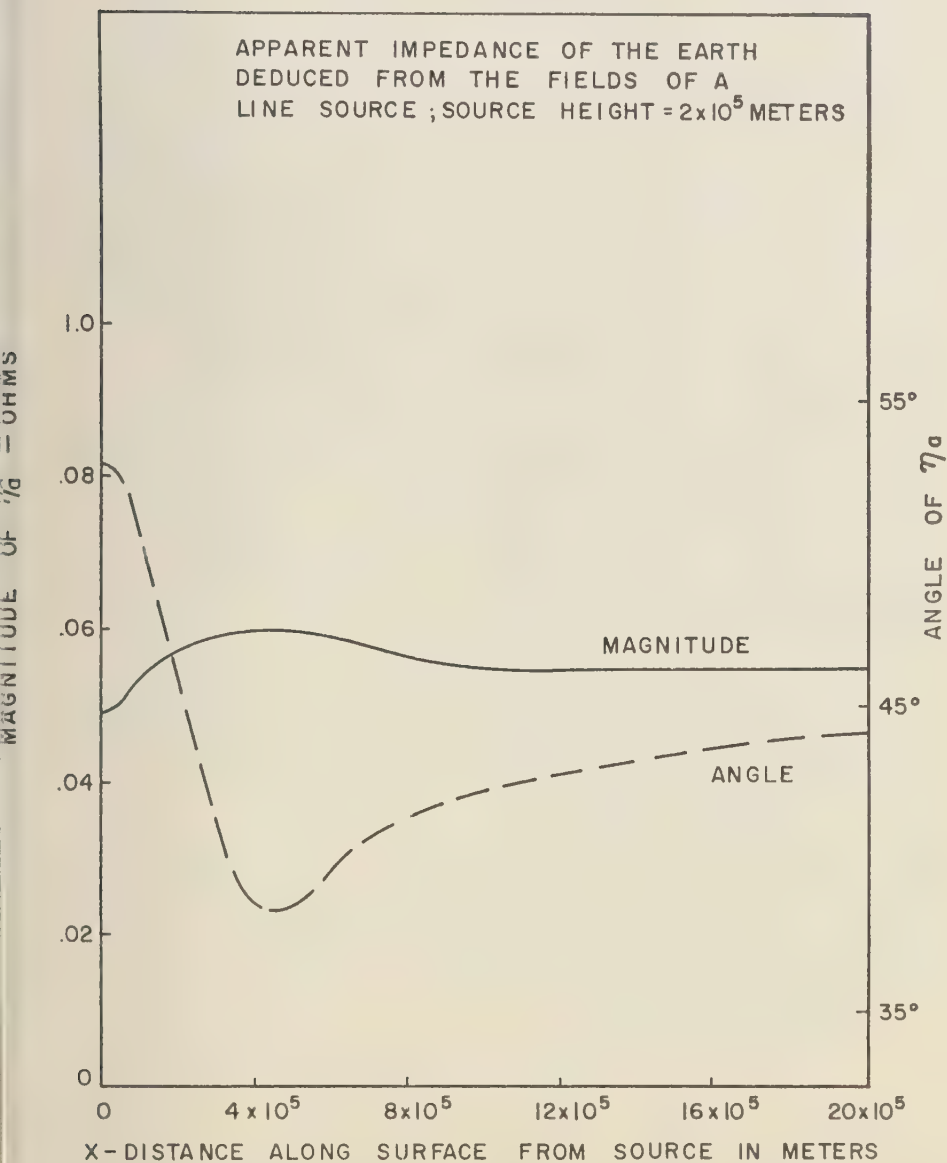


Fig. 6. Magnitude and angle of apparent impedance of the earth, $\eta_a = \bar{E}_z / \bar{H}_z$.

can be deduced from plane wave incidence. The maximum variation of the values calculated in the case considered here from the plane wave value is about 10 per cent in amplitude and 8° in phase angle. The reason that this deviation is so small is considered to be that the curvature of the wave fronts within the earth is much less than the curvature of the incident fronts, owing to the velocity of propagation ratio of about 5000 to 1 (for the problem parameters chosen) in favor of free space over the earth. For a more realistic earth model it is expected that apparent impedance calculations based on plane wave assumptions would differ by a maximum of about 10 to 20 per cent from results obtained by assuming the earth to be in the near field of the incident radiation.

The numerical results obtained for the values of the field components at the earth's surface are dependent on the choice of values for the parameters h , σ_e , and ω . Exact field values for parameter choices other than those taken herein may be deduced from equation 14 and the associated magnetic field relation. However, it may be noted in general that lowering of radiator height with other parameters held constant increases the values of field components at the earth's surface. By an application of rules of scaling one can deduce other values of fields; for instance, if the radiator height were halved (100 km), the earth conductivity doubled (1/4000 mho/meter), and the frequency doubled (0.6 radian/second), the fields at the earth's surface would be given by Figures 3, 4, and 5 with their ordinates doubled and their abscissas halved.

APPENDIX

Validity of Approximation. In order to evaluate the integral

$$I(x, y) = \int_{-\infty}^{\infty} \frac{2C}{C + \sqrt{S_0^2 - S^2}} \cdot e^{-jk[Sx + C(y+h)]} \frac{dS}{C} \quad (15)$$

the approximation

$$C + \sqrt{1 - S^2} \cong \begin{cases} jS, & S < 0 \\ -jS, & S > 0 \end{cases}$$

was made (see equation 12 et seq. in text). For values of $S \geq A$, where $A \gg 1$ this approxima-

tion appears reasonable as one notes that $|C|$ asymptotically approaches $|S|$ for increasing values of $|S|$. For values of $|S| < A$, however, the validity of this approximation is not obvious. For this reason it is necessary to evaluate the true value of the integral on $[0, A]$ as carefully as possible and compare the result with the result obtained on $[0, A]$ with the approximation indicated above. This has been done in the analysis that follows. The evaluations of these two integrals (one with the true integrand and one with the approximate integrand) require approximations that are considered acceptable.

The integral (1) may be written as

$$\begin{aligned} I &= \int_{-\infty}^0 \frac{2C}{C + \sqrt{S_0^2 - S^2}} \cdot e^{-jk[Sx + C(y+h)]} \frac{dS}{C} + \int_0^{\infty} (\text{same}) \\ &= \int_0^{\infty} \frac{2C}{C + \sqrt{S_0^2 - S^2}} \cdot e^{-jkC(y+h)} [e^{-jkSx} + e^{jkSx}] \frac{dS}{C} \quad (16) \end{aligned}$$

where the change of variable $S = -S$, $C = C$ has been made in the first member before combination under the single integral sign. Now this integral may be written as

$$I = \int_0^{\infty} = \int_0^A + \int_A^{\infty} \quad (17)$$

where A is a value of S that may be chosen to satisfy several criteria:

$$A \gg 1$$

$$A \ll |S_0| \quad A \ll 1/k\rho_2$$

In the \int_A^{∞} the approximation

$$C \cong \begin{cases} jS, & S < 0 \\ -jS, & S > 0 \end{cases}$$

introduces small error since $A \gg 1$. The true integral \int_0^A will be evaluated and compared with the value resulting under the approximation to yield an estimate of the error in the approximation on the range 0 to A .

$$\int_0^A (\text{true}) = 2 \int_0^A \frac{2C}{C + \sqrt{S_0^2 - S^2}} \cdot e^{-jkC(y+h)} \cos kSx \frac{dS}{C} \quad (18)$$

but $kSx \ll 1$ and $|S_0| \gg S$, $|S_0| \gg C$, so that

$$\int_0^A (\text{true}) \cong 4 \int_0^A \frac{e^{-jkC(y+h)}}{S_0} dS \quad (19)$$

also $kC(y+h) \ll 1$; consequently

$$\begin{aligned} \int_0^A (\text{true}) &\cong \int_0^A \frac{4}{S_0} [1 - jkC(y+h)] dS \\ &= \frac{4}{S_0} \int_0^A dS - \frac{4jk(y+h)}{S_0} \int_0^{\sin \theta - A} \cos^2 \theta d\theta \\ &= \frac{4}{S_0} \left\{ A - \frac{jk(y+h)}{2} A \sqrt{1 - A^2} \right\} \\ &= \frac{4A}{S_0} - 0 \left\{ \frac{jk(y+h)A}{2S_0} \right\} \quad (20) \end{aligned}$$

Now evaluation with the approximations

$$C \cong \begin{cases} jS, & S < 0 \\ -jS, & S > 0 \end{cases}$$

as follows:

$$\begin{aligned} \int_0^A (\text{appx}) &= \int_0^A \frac{2}{-jS + \sqrt{S_0^2 - S^2}} \cdot e^{-jk(-jS)(y+h)} \cos kSx dS \\ &\cong \frac{4}{S_0} \int_0^A e^{-kS(y+h)} dS \\ &= \frac{4}{S_0} \left[\frac{1}{k(y+h)} - \frac{1}{k(y+h)} e^{-kA(y+h)} \right] \\ &= \frac{4}{kS_0(y+h)} [1 - e^{-kA(y+h)}] \end{aligned}$$

$$\begin{aligned} &= \frac{4}{kS_0(y+h)} \left[1 - 1 + kA(y+h) \right. \\ &\quad \left. - \frac{[kA(y+h)]^2}{2!} + \text{smaller terms} \right] \\ &= \frac{4A}{S_0} - 0 \left[\frac{4kA(y+h)}{S_0(2!)} \right] \quad (21) \end{aligned}$$

So that it is seen the approximation introduces an error of magnitude

$$\frac{\sqrt{2} 4k(y+h)}{2!} \cdot 100\% \cong 0.1\%$$

REFERENCES

- Benkova, N. P., Magnetic storms and systems of electric currents, *Trudy Nauch.-Issledovatel. Inst. Zemnogo Magnetizma*, no. 10, Leningrad, 1953, Translated by Air Technical Intelligence Center, Wright-Patterson Air Force Base, Ohio, ASTIA Report AD 162 841.
- Erdelyi, A., W. Magnus, F. Oberhettinger, and F. G. Tricomi, *Higher Transcendental Functions*, vol. 2, Bateman Manuscript Project, McGraw-Hill Book Co., 1953.
- Erdelyi, A., W. Magnus, F. Oberhettinger, and F. G. Tricomi, *Tables of Integral Transforms*, vol. 1, Bateman Manuscript Project, McGraw-Hill Book Co., 1954.
- Hasegawa, M., On the position of the focus of the geomagnetic S_q current system, *J. Geophys. Research*, 65, 1437-1447, 1960.
- Jacobs, J. A., and K. Sinno, World wide characteristics of geomagnetic micropulsations, *Univ. Brit. Columbia Geophys. Lab. Rept. 2*, April 1960.
- Law, P. F., and B. M. Fannin, Radiation from a current filament above a homogeneous earth with application to micropulsations, *Univ. Texas Elec. Eng. Research Lab. Rept. 118*, November 1960.
- Tamm, I. E., Currents in the ionosphere causing variations in the earth's magnetic field, *Izvest. Akad. Nauk SSSR*, 8 (2), 1944. Translated by Air Technical Intelligence Center, Wright-Patterson Air Force Base, Ohio, ASTIA Report AD 37 552.

(Manuscript received November 28, 1960;
revised January 26, 1961.)

Ionospheric Electron Content and Its Variations Deduced from Satellite Observations

K. C. YEH

Department of Electrical Engineering

AND

G. W. SWENSON, JR.

*Departments of Electrical Engineering and Astronomy
University of Illinois, Urbana, Illinois*

Abstract. A procedure is given for correcting the effect of refraction and the high-frequency approximation when Faraday rotation measurement is available on two frequencies. This method is used to analyze the records obtained on the ground of radio transmissions from satellite 1958 δ_2 (Sputnik III). The result reveals strong diurnal as well as anomalous seasonal variations. The depression in electron content during magnetic storms is identified, and the preliminary observational result is not inconsistent with the drift theory.

Introduction. Since the advent of artificial satellites it has been observed that their radio signals exhibit strong ionospheric influences. One of the easily recognizable effects is produced by the rotation of the plane of polarization of the wave along the ray, known as Faraday rotation. For a linearly polarized transmitting signal the amplitude record of a linearly polarized receiving antenna approximates a rectified sine wave. The use of this phenomenon has already produced valuable information about the behavior of the ionosphere [Garriott, 1960; Little and Lawrence, 1960; Hame and Stuart, 1960]. This paper is concerned with the observation of the same phenomenon, from which the electron content in a column of unit cross-sectional area is obtained. All the results given here are based on the analysis of satellite 1958 δ_2 (Sputnik III) signals on 20 and 40 Mc/s. The data were obtained during a period of more than a year, from September 1958 to December 1959.

Theory. The rotation of the plane of polarization of a linearly polarized wave is given by the familiar relation, first used in connection with moon echo investigations [Browne, Evans, and Hargreaves, 1956],

$$\Omega = \frac{K_1}{f^2} \int NM \, dh \quad \text{radians} \quad (1)$$

where

$K_1 = 2.97 \times 10^{-2}$ in mks rationalized units.

f = radio frequency in c/s.

N = electron volume density.

$M = H \cos \phi \sec i$.

H = earth's magnetic field strength.

ϕ = angle between the ray and the earth's field.

i = vertical angle of the ray.

dh = differential height element.

In addition to the quasi-longitudinal approximation in deriving (1), it implicitly assumed that: (a) the frequency is so high that refraction can be neglected; (b) frequency is much higher than the maximum plasma frequency; and (c) magnetoionic path splitting does not occur.

Of these three conditions the first two are most serious on 20 Mc/s. The effect of splitting of the magnetoionic components is negligible if the vertical angle is less than about 40° [Garriott, 1960]. On the other hand, the plasma frequency may rise to more than half the radio frequency, resulting in considerable refraction for an oblique ray. Therefore, steps must be taken to correct for these effects. A formula taking the first two

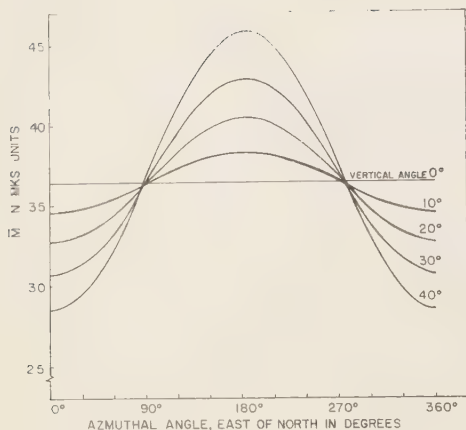


Fig. 1. The variation of \bar{M} at the University of Illinois.

effects into account and correct to the second order is [Yeh, 1960]¹

$$\Omega = \frac{K_1}{f^2} \left\{ \int N M dh + \frac{1}{2} \sec^2 i_0 \frac{K_2}{f^2} \int N^2 M dh \right\} \quad (2)$$

where $K_2 = 80.62$ and $i_0 =$ vertical angle of the ray at the bottom of the ionosphere.

The function M is a geometric factor which involves the earth's magnetic field. Its properties have been studied extensively using spherical harmonics [Yeh and Gonzalez, 1960]. The form of (1) and (2) is inconvenient because it involves integration of the product of N and M . Since M varies slowly, it is permissible to take M outside of the integral, the actual value depending on the profile of the ionization density and the direction of the ray. It has been found empirically that the value of M at 350 km corresponds closely with the desired average. Let us denote this average value by \bar{M} . Equation 2 can be written

¹ One of us (K. C. Y.) is grateful to Dr. J. E. Titheridge for pointing out the missing factor $\frac{1}{2}$ in equation 3 of the given reference. The correct form of (3) is

$$\sec i = \sec i_0 (1 + \frac{1}{2} \times \tan^2 i_0)$$

All the equations following would have to be corrected accordingly except (8) and (11) which are most useful to obtain the electron content.

$$\Omega = \frac{K_1 \bar{M}}{f^2} \left\{ \int N dh + \frac{1}{2} \sec^2 i_0 \frac{K_2}{f^2} \int N^2 dh \right\} \quad (3)$$

The dependence of \bar{M} on the azimuthal and vertical angle for a station at the University of Illinois is shown graphically in Figure 1. It is calculated from the published Gaussian coefficients [Finch and Leaton, 1957].

If Faraday rotation measurements are obtained on two frequencies f and mf , the application of (3) yields

$$\int N dh = \frac{f^2}{K_1 \bar{M}} \frac{1}{m^2 - 1} \cdot \{ m^4 \Omega(mf) - \Omega(f) \} \quad (4)$$

Equation 4 gives the electron content correct to the second order.

Method of analysis. In practice the number of rotations along the ray is not obtained experimentally. Instead, the measurable quantity is the differential rotation as the satellite moves across the sky. Consequently, it is convenient to make further simplifying assumptions in order to facilitate the analysis.

1. High-pass approximation: It is assumed that the satellite is practically outside the ionosphere, so that the vertical motion of the satellite is unimportant. The lower limit of the height is arbitrarily chosen as 1000 km, below which lies almost all the ionosphere. When this condition is not met the satellite records can still be analyzed but with reduced accuracy. It is hoped that the result of analysis of these low-pass records will be published in a later communication.

2. Negligible horizontal gradients: It is assumed that the ionosphere is spherically stratified. Several satellite records have been subdivided into two or three portions and analyzed separately. The values obtained from different portions of the same record agree within about 5 per cent. Strong horizontal gradients are known to exist during sunrise hours. It is expected that the electron content obtained during these hours will be less accurate.

With these two assumptions we see from (3) that the differential rotation of polarization is entirely produced by the change in geometry. The orbit of the satellite being known, the

analysis is reduced to a problem of geometry. Typically, a portion of a record 1 or 2 minutes long is chosen when the satellite is near its point of closest approach. The number of polarization rotations during the interval is counted. Since the geometric factor \bar{M} has already been calculated (see Fig. 1), the integrated electron density can be computed very easily. If measurements are available on two frequencies (e.g., 10 and 40 Mc/s) the electron content correct to the second order can be obtained.

A study has been made of the correction factor needed to obtain the electron content correct to the second order. Let $(\int N dh)_1$ denote the electron content obtained by using the first-order theory, and $(\int N dh)_2$ the electron content obtained by using the second-order theory. Equation 3 can then be written as

$$\left(\int N dh \right)_2 / \left(\int N dh \right)_1 = 1 - \mu \quad (5)$$

where μ is the correction factor and is defined by

$$\mu = \frac{K_2}{2f^2} \sec^2 i_0 \frac{\int N^2 dh}{\int N dh} \quad (6)$$

If it is assumed that the shape of the ionization density profile is constant but that the absolute magnitude varies, the ratio of the integrals in (6) is then proportional to the electron content. Therefore, a plot of the correction factor versus the product of $\sec^2 i_0$ and the electron content is made; it is shown in Figure 2. A line is drawn to indicate the average behavior. It will be noted that, at low values of the abscissa, the observed correction factors are consistently less than indicated by the 'average line.' The observations were made near sunrise, when it is to be expected that large horizontal gradients will exist, with higher electron densities to the southeast. This, combined with the magnetic field geometry, could cause the correction factor to be less than that calculated on the assumption of spherical symmetry of N . Low values of the abscissa also emphasize the errors in the correction factor which result from erroneous estimates of the satellite spin rate, as discussed below. Although the points are scattered, their spread is less than 10 per cent from the average line. When, as often happens, the satellite signal on 40 Mc/s is too

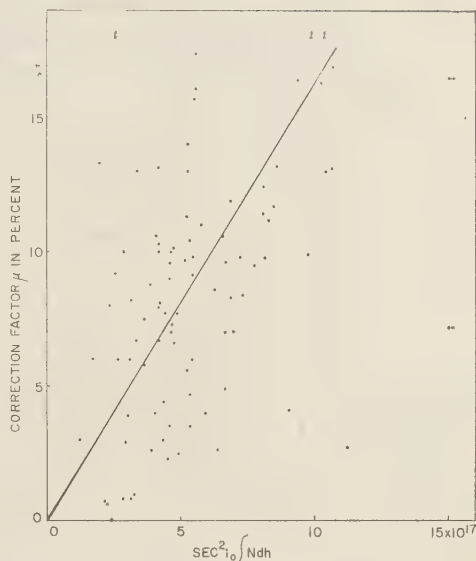


Fig. 2. The dependence of the correction factor.

weak to record, the electron content deduced from the first-order theory is corrected according to the average line of Figure 2.

Interestingly, if it is assumed that the profile takes the form of a Chapman function,

$$N = N_0 \exp \frac{1}{2}(1 - z - e^{-z}) \quad (7)$$

the line in Figure 2 gives a scale height of 98 km. This value is consistent with those used by several authors [Nicolet, 1954; Wright, 1960].

During the early life of satellite 1958 δ_2 , many observed polarization fading on the record can only be explained as due to the spin motion of the satellite. As the satellite spins, the dipole antenna generates a cone. If the direction of the observer is outside the cone the Faraday fading is expected to be modulated by the spin motion but it will not alter the average Faraday effect [Thomson, 1958]. If the direction of the observer is inside the cone, however, the observed fading is the algebraic sum of the Faraday and the spin fading. The algebraic sign can be determined by observing the fading on two frequencies to see whether the inverse-frequency-squared law given by (1) is approximately satisfied. Sometimes, the Faraday fading and the satellite spin fading may have comparable magnitudes, making accurate measurements of electron content difficult.

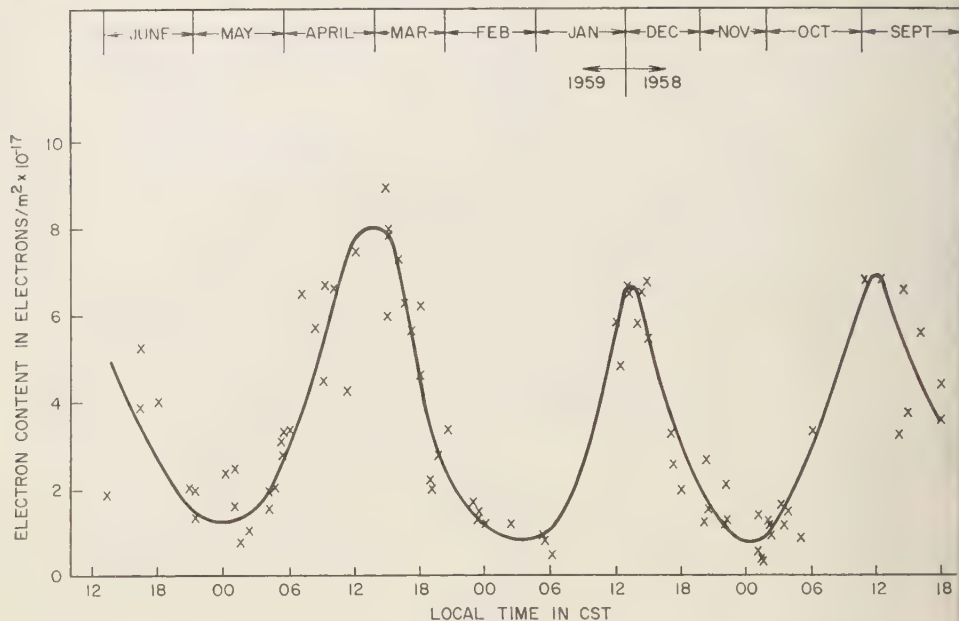


Fig. 3. Hourly variation of electron content deduced from south-bound passages.

Experimental results. The electron contents deduced from observation of satellite 1958 δ_2 signals on 20 and 40 Mc/s are shown in Figures 3 and 4. Figure 3 gives the high-pass observations of satellite 1958 δ_2 while it is descending from

northwest to southeast, and Figure 4 is for ascending high passages going from southwest to northeast. Owing to the westward motion of the nodes the time of observation advances each day. This enables us to obtain the diurnal varia-

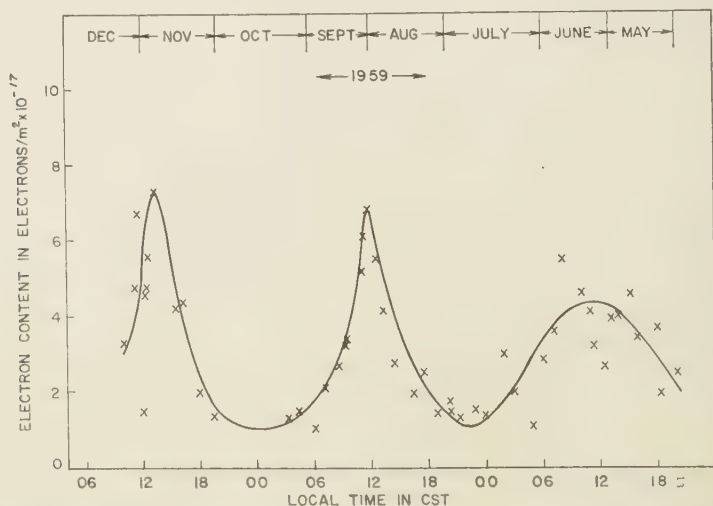


Fig. 4. Hourly variation of electron content deduced from north-bound passages.

tion of the electron content; it shows up clearly in Figures 3 and 4. The electron contents measured on magnetically disturbed days have been deleted from these figures in an effort to show the quiet-day behavior, since it is discovered that the values measured on disturbed days may depart considerably from the average behavior.

During the period of interest the sunspot number is found to vary over a wide range. Empirically, the critical frequency of the F_2 region is related to the sunspot number through the following relation [Allen, 1948]:

$$(f_c)^2 \propto (1 + 0.02R) \quad (8)$$

where R is the mean Zurich sunspot number. The fact that the electron density is proportional to the sunspot number is consistent with the present theory of ionospheric formation [Ratcliffe, 1959]. Since the equation of continuity that governs the electron density is linear, the electron content would be expected also to follow a law given by (8). Hence the measured electron content shown in Figures 3 and 4 can be normalized through (8) to give the variation of electron content for zero sunspot number. It is believed that superimposed on the constant-background

ultraviolet brightness of the permanent corona there are variable, long-lived bright coronal patches whose intensity is controlled, with a delay of a few days, by the underlying spot groups [Kiepenheuer, 1946]. Since the effect we are seeking is seasonal, this delay is unimportant to us.

The result of this normalization is shown in Figures 5 and 6. We see immediately that although the nighttime electron content shows no seasonal variation the daytime maximum certainly exhibits the seasonal anomaly. It is known that the maximum critical frequency of the F_2 region has higher values during the winter than during the summer [Ratcliffe and Weekes, 1960]. The present result indicates that the same is true for total electron content, though perhaps to a lesser degree.

Many magnetic storms have occurred during the period of the observations. During the few days following a sudden commencement the electron content is found to be consistently lower than the quiet-day average. This shows up very clearly in Table 1, where the depressions of the electron content measured during seven magnetic storms are listed. These depressions are

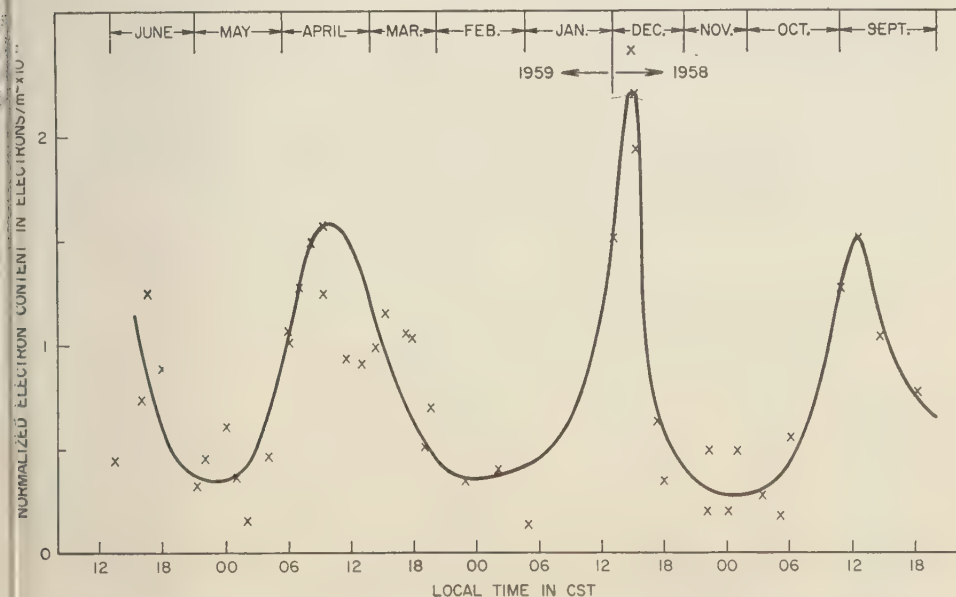


Fig. 5. Hourly variation of electron content normalized to zero sunspot number (south-bound passages).

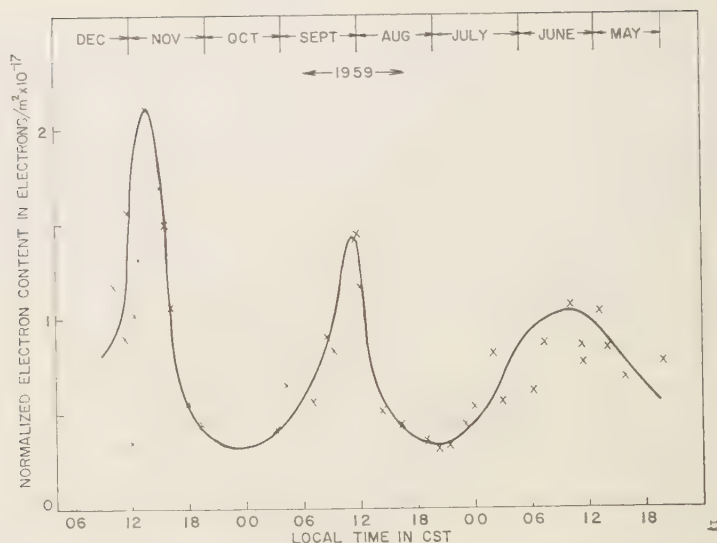


Fig. 6. Hourly variation of electron content normalized to zero sunspot number (north-bound passages).

believed to be real, since the standard deviation of the quiet-day electron content from the average diurnal variation is 0.84×10^{17} electrons/m² (shown in Figs. 3 and 4) and is 3.51×10^{17} electrons/m² during disturbed days. Thus the storm effect is more than 4 times the background fluctuations.

The drift theory seems to be quite successful in explaining the variation of electron density during magnetic storms [Maeda and Sato, 1959]. The University of Illinois is situated at a latitude where a depression in F_2 critical frequencies is

most often observed during storms. According to the theory the storm effect is maximum around noon and minimum at night. It is interesting that the only nighttime value in Table 1 (December 4, 1958) gives the smallest depression.

Another interesting observation is that large depressions of electron content have been observed even during relatively quiet days (Yeh, to be published). It seems that such depressions cannot be explained as a direct cause of the change in the solar activity. This type of disturbance behaves somewhat like those accompanying magnetic storms in that it may last a few days and cover a region of at least 400 km. The cause of this quiet-day ionospheric disturbance is unknown.

Conclusion. The observation of Faraday rotation on 20 Mc/s should be corrected for refraction and the high-frequency approximation. A correction procedure is given here, usable when such measurements are available on two frequencies. A study of the correction factor yields a scale height of 98 km, which is very close to the value used by others. This fact increases our confidence in the method.

The result of the analysis indicates a strong diurnal variation in electron content. The anomalous seasonal variation is obtained after

TABLE 1. The Depression of Electron Content during Magnetic Storms

Date	Local Time	$\Delta \int N dh$ electrons/m ²
Sept. 25, 1958	1144	-2.9×10^{17}
Sept. 26, 1958	1202	-4.3
Dec. 4, 1958	2009	-0.8
Mar. 26, 1959	1434	-4.4
Mar. 27, 1959	1541	-6.0
Apr. 10, 1959	1004	-3.5
Apr. 10, 1959	1148	-3.3
May 24, 1959	1435	-0.9
Sept. 4, 1959	1204	-3.5
Sept. 5, 1959	1050	-4.5
Sept. 6, 1959	1114	-3.0
Nov. 28, 1959	1240	-2.8

the measured values are normalized to a constant ionizing intensity. It seems that the loss of electrons during summer more than balances the increased production rate, resulting in a lower density in summer than in winter.

Two kinds of ionospheric storms have been observed, one accompanied and the other not accompanied by magnetic storms. The behavior of the electron content during magnetic storms is consistent with the drift theory.

Acknowledgment. Many people have assisted in carrying out the investigation reported in this paper. J. P. McClure, J. G. Card, W. W. Cochran, and J. E. Bowers collected data and constructed and maintained the field station. W. G. Johnston, H. Chow, and I. Ivanoff helped in the analysis of records. V. H. Gonzalez programmed the geometric factor M on Illiac and computed the curves shown in Figure 1. The orbital information was provided by the Smithsonian Astrophysical Observatory and the National Aeronautics and Space Administration. This project was initiated under grant Y/32.40/266 from the U. S. National Committee, IGY (National Science Foundation), and has been continued since July 1959 under grant NsG2459 from the National Aeronautics and Space Administration.

REFERENCES

Allen, C. W., Critical frequencies, sunspots, and the sun's ultraviolet radiation, *J. Geophys. Research*, **53**, 433-448, 1948.
 Browne, I. C., J. E. Evans, and J. K. Hargreaves, Radio echoes from the moon, *Proc. Phys. Soc. London, B*, **69**, 901, 1956.
 Finch, H. F., and B. R. Leaton, The earth's main magnetic field—epoch 1955.0, *Monthly Notices*

Roy. Astron. Soc., Geophys. Suppl., **1**, 314, 1957.
 Garriott, Owen K., The determination of ionospheric electron content and distribution from satellite observations, I and II, *J. Geophys. Research*, **65**, 1139-1157, 1960.
 Hame, T. G., and W. D. Stuart, The electron content and distribution in the ionosphere, *Proc. IRE*, **48**, 1786-1787, 1960.
 Kiepenheuer, K. O., On the relations between ionosphere, sunspots, and solar corona, *Monthly Notices Roy. Astron. Soc.*, **106**, 515, 1946.
 Little, C. G., and R. S. Lawrence, The use of polarization fading of satellite signals to study the electron content and irregularities in the ionosphere, *J. Research NBS*, **64D**, 335-346, 1960.
 Maeda, K. I., and T. Sato, The F region during magnetic storms, *Proc. IRE*, **47**, 232-239, 1959.
 Nicolet, M., Dynamic effects in the high atmosphere, in *The Earth as a Planet*, edited by G. P. Kuiper, University of Chicago Press, pp. 644-712, 1954.
 Ratcliffe, J. A., Recent trends in the theory of the ionosphere, Presidential address, Year Book of the Physical Society, 1959.
 Ratcliffe, J. A. and K. Weekes, The ionosphere, in *Physics of the Upper Atmosphere*, edited by J. A. Ratcliffe, Academic Press, New York, pp. 378-456, 1960.
 Thomson, J. H., The rotation of the first Russian earth-satellite, *Phil. Mag.*, **3**, 912, 1958.
 Wright, J. W., A model of the F region above $h_{\max} F_2$, *J. Geophys. Research*, **65**, 185-191, 1960.
 Yeh, K. C., Second-order Faraday rotation formulas, *J. Geophys. Research*, **65**, 2548-2550, 1960.
 Yeh, K. C., and V. H. Gonzales, Note on the geometry of the earth magnetic field useful to Faraday effect experiments, *J. Geophys. Research*, **65**, 3209-3214, 1960.

(Manuscript received December 12, 1960;
 revised January 16, 1961.)

A Note on 106.1-Mc Auroral Echoes Detected at Stanford Following the Solar Event of November 12, 1960

R. L. LEADABRAND, W. E. JAYE, AND R. B. DYCE

*Stanford Research Institute
Menlo Park, California*

Abstract. Auroral echoes detected following the solar disturbance of November 12, 1960, were obtained at Stanford using a 106.1-Mc radar. These echoes allowed the transit time of the auroral particles from the solar disturbance to be determined (20 hr, 5 min). A particularly rapidly moving echo (100 km/s) was detected during the display.

Introduction. An unprecedented solar disturbance occurred on November 12, 1960, at 1325 GMT. A polar-cap event with 15-db absorption occurred almost immediately following the solar event and lasted 24 hours (Leinbach, private communication, 1960). It was predicted that such a major solar event would cause a sizable auroral display to occur. As during major auroral disturbances, an extensive aurora was anticipated far south of the normal auroral zone. (Subsequently a Special World Interval was declared.)

In order to examine the intensity and extent of the aurora that was expected, a 106.1-Mc radar located at Stanford was operated during the night of November 12, 1960.

It has been reasonably well established that auroral echoes arise most frequently when the line-of-sight from the radar intersects the earth's magnetic field lines at right angles. The scattering polar diagram of the magnetic field aligned scatterers is such that at 100 Mc several degrees of off-perpendicular angle will still allow auroral reflections to be obtained.

The location of the Stanford site is such that perpendicular reflection from the earth's magnetic field lines can be obtained for scattering volumes within the *E* and *F* regions. The *E*-region perpendicular-intersection locus is located quite far south, and the *F*-region perpendicular-intersection locus is located almost over the United States-Canadian border. The magnetic field auroral geometry for the Stanford site is shown in Figure 1, for 100, 200, and 300 km.

Thus, at the Stanford location it is possible to study auroral reflections over a fairly large range

of altitude and height [Schlobohm, Leadabrand, Dyce, Dolphin and Berg, 1959; Leadabrand and Peterson, 1958], although in general aurora seldom occurs far enough south with sufficient intensity to permit the 106-Mc radar to obtain reflections. The Stanford site is unique inasmuch as *F*-region auroral echoes are usually obtained more commonly than *E*-region echoes during intense displays.

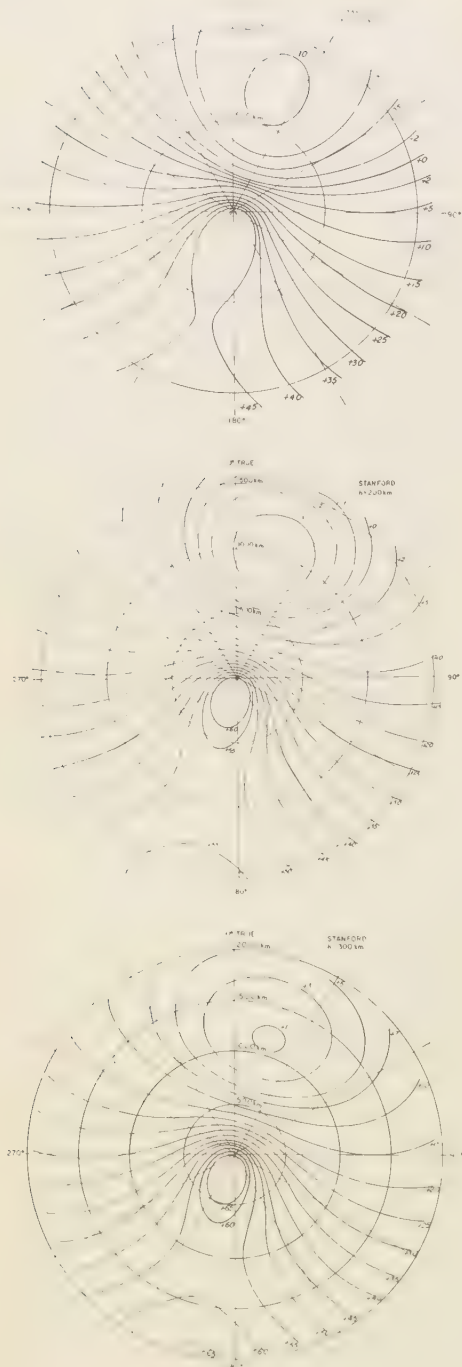
Equipment. The radar antenna used for these experiments consisted of a 60-foot parabolic reflector, steerable in elevation and azimuth. Vertically polarized signals are transmitted, and vertically and horizontally polarized signals are received. The system parameters used for this experiment are listed in Table 1.

TABLE 1

Frequency, Mc	106.1
Peak power, kw	40
Antenna gain (one way), db	24
Pulse length, ms	1
PRF, cps	75
Maximum unambiguous range, km	2000
Receiver bandwidth, kc	6
Minimum detectable signal, dbm	-115

The antenna was scanned (at the discretion of the operator) in elevation and azimuth throughout the test. Magnetic tape recordings of the detected signals were made for later analysis.

Description of results. The radar was placed in operation at approximately 2100 local time. No auroral echoes were seen until about 2233 local time when a weak return was detected in the magnetic north direction at 1400-km range. The return was at maximum amplitude with the



antenna at zero degrees elevation. This auroral echo disappeared after about 5 minutes.

At 2400 local time a local FM broadcast station left the air thereby eliminating about 10 db of interference noise level from the radar. At this time weak returns were detected from the extremities of the 100-km orthogonality locus (true azimuths of 310° and 75° at ranges of 1000 km).

By moving the antenna in azimuth and elevation, echoes were detected all along the 100-km height orthogonality locus between azimuths of 310° and 75° true. This condition remained fairly constant with the echoes lying along this locus, increasing and decreasing somewhat in amplitude with periods of some 15 minutes. In general, only returns on the vertically polarized receiver channel were seen.

At about 0130 local time, echoes started to appear from magnetic north at great range (1000 to 1400 km). These echoes were unusually strong (30 db S/N) and maximized at an elevation angle of about 10° . Figure 2 indicates that for ranges of 1000–1400 km and elevation angle of 10° the auroral scatterers were located in the F layer at about 300 km height.

Primarily this echo was seen in the horizontal polarization channel only (orthogonal to the transmitted energy). In general, past experiments have indicated that the polarization of the transmitted energy is preserved [Presnell, Leadabrand, Peterson, Dyce, Schlobohm, and Berg, 1959]. Faraday rotation in the lower F region is believed to be responsible for this observation as the radio wave was close to, but not exactly at, right angles to the magnetic field all along the propagation path.

Returns at ranges of 800 to 1600 km continued to be seen in the magnetic north direction for about 30 minutes after first being detected. After that period only echoes lying along the 100-km height remained.

At about 0408:10 local time a very unusual auroral echo was seen. The antenna was pointing at 2° elevation at an azimuth of about 71° true. Suddenly a strong auroral echo appeared (at a range of 800 km) and rapidly moved in range to about 1200 km in a period of about 10 seconds.

←

Fig. 1. Magnetic field geometry for Stanford (100-, 200-, and 300-km heights).

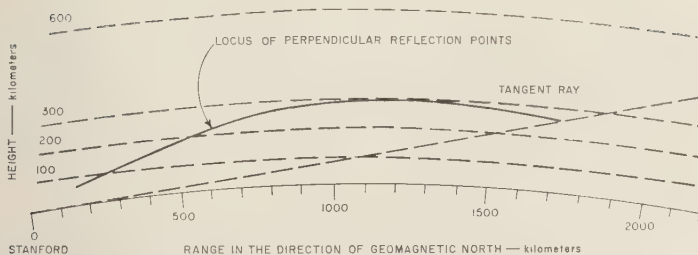


Fig. 2. Magnetic field geometry for meridian plane at Stanford.

This phenomenon repeated itself three more times (see Fig. 3 for a range-time display of these echoes).

The apparent velocity indicated by a portion of the range change observed was a maximum of 100 km/s. The average velocity was of the order of 40 km/s. Such velocities correspond very nearly to the east-west speeds observed visually in auroral arcs at 'break-up.'

The *E*-region auroral echoes continued to be seen until about 0600 PST. The radar was shut down at 0700 PST.

Discussion of results. The results described in this paper are believed to be of considerable interest to those trying to understand auroral phenomena as related to solar disturbances. The use of the 106-Mc radar, which is located quite far south of the normal auroral zone, allows the determination of the occurrence time of the solar

disturbance-produced aurora. In this case a time lapse between the solar disturbance and the aurora was about 20 hours and 5 minutes.

The real utility of radar in the study of aurora arises from its ability to determine the position in space over which auroral echoes are seen. For example, during this test aurora could be shown to exist in the *E* region quite close to Stanford and in the *F* region somewhat close to the United States-Canadian border.

The unusually rapid motion of the auroral echoes obtained during this test has never before been reported in the literature. It is interesting to note that the velocity represented by these range changes would cause a Doppler shift in excess of 30 kc if the range change represented a true velocity. Such a high Doppler shift would, however, place the entire return outside the receiver 6-kc bandwidth. Therefore, these range

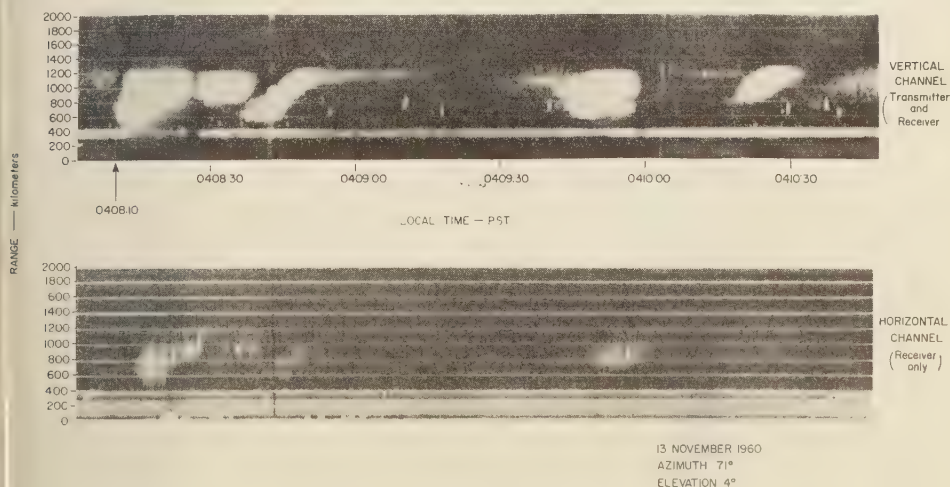


Fig. 3. Range-time display of rapidly moving auroral echoes.

changes cannot be interpreted in terms of actual motion as the returned energy was seen within the receiver bandwidth.

Such a high velocity represented by this range change is rarely seen except when a simple auroral arc commences to break up. Then an apparent motion is detected in the rayed structure that runs from east to west. The apparent velocity represented by such motion is of the same order of magnitude as the range change observed in the auroral echoes described here.

An alternative explanation would be to imagine a 'fire hose' of primary auroral particles being injected into the earth's atmosphere with the nozzle of the fire hose veering rapidly. The only problem with this interpretation is that such a phenomenon would cause high Doppler shifts.

Summary. Auroral echoes associated with the solar event were obtained at Stanford using a 106-Mc radar. Several interesting facts were obtained:

1. The major onset of the aurora associated with the November 12, 1960, solar disturbance occurred approximately 20 hours and 5 minutes afterward.

2. The aurora occurred predominantly in the *E* region not far to the north of Stanford and persisted throughout the night (2400 until 0600 local time).

3. At 0130 local time *F*-region auroral echoes which were of orthogonal polarization (due to

Faraday rotation) to the transmitted polarization were observed to last for approximately 30 minutes.

4. Suddenly occurring auroral echoes were seen at 1209 GMT, and these echoes moved in range extremely rapidly. The apparent velocity represented by the range changes were as great as 100 km/s.

5. The rapidly moving echoes are interpreted as a result of a sudden break-up of auroral forms; the apparent velocity represented by the range change for these echoes is comparable to the east-west motion observed during the break-up of a simple auroral arc.

REFERENCES

- Leadabrand, R. L., and A. M. Peterson, Radio echoes from auroral ionization detected at relatively low geomagnetic latitudes, *Proc. IRE: PGAP, AP-6*, 65-79, 1958.
- Presnell, R. I., R. L. Leadabrand, A. M. Peterson, R. B. Dyce, J. C. Schlobohm, and M. R. Berg, VHF and UHF Radar Observations of the aurora at College, Alaska, *J. Geophys. Research*, *64*, 1179-1190, 1959.
- Schlobohm, J. C., R. L. Leadabrand, R. B. Dyce, L. T. Dolphin, and M. R. Berg, High-altitude radio echoes from auroral ionization detected at a geomagnetic latitude of 43°, *J. Geophys. Research*, *64*, 1191-1196, 1959.

(Manuscript received January 16, 1961;
revised January 30, 1961.)

A Local Reduction of *F*-Region Ionization Due to Missile Transit

HENRY G. BOOKER

*Center for Radiophysics and Space Research
Cornell University, Ithaca, New York*

Abstract. This paper describes and interprets an unusual echo received by local ionospheric sounders for a period of about one-half hour subsequent to the firing of Vanguard II.

Introduction. Vanguard II was fired from Cape Canaveral shortly before 1100 EST on February 17, 1959. The behavior of the ionosphere was monitored by the U. S. Army Signal Radio Propagation Agency using an ionospheric sounder on Cape Canaveral itself and a second ionospheric sounder on Grand Bahama Island. Grand Bahama Island is about 300-km down-range and is closer than Cape Canaveral to the location where the missile penetrated the *F* region. The geometry is shown approximately in Figure 1; the penetration point indicated is judged from the ionograms themselves. Ionograms were taken on Grand Bahama Island at frequent intervals, and four sketches are shown

in Figure 2. Qualitatively similar, but less well-developed, phenomena were observed by the sounder on Cape Canaveral. Ionograms of the type shown in Figure 2 have been obtained in association with a number of missile firings.

In Figure 2 the letter *F* indicates a regular *F*-region echo whereas the letter *M* indicates an unusual echo associated with passage of the missile. At 1102½, 1109, and 1115½ EST a first, and even a second, multiple of the missile-produced echo can be seen. After 1115½ EST the primary echo associated with the missile disappeared. The first and second multiples could still be seen, however, as shown in Figure 2*d*. The first multiple disappeared at 1130 EST, and

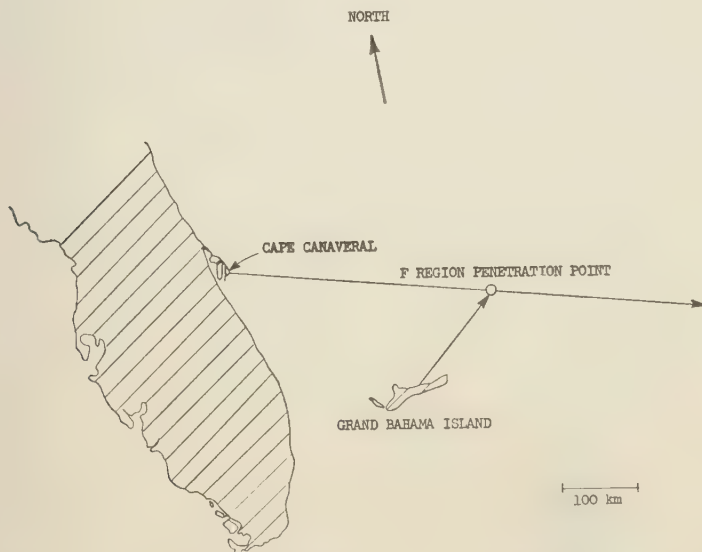


Fig. 1. Indicating the approximate geometry involved in firing Vanguard II.

shortly thereafter unusual echoes could no longer be satisfactorily identified.

Figure 2 shows that the penetration-frequency for the unusual echo is substantially identical with the penetration-frequency of the layer. It is also noticeable that the minimum frequency

for the unusual echo is clearly identified by means of a loop in the trace. During the first 3 minutes this minimum frequency decreased from about 10 Mc/s to 7.5 Mc/s. It then remained steady for about 10 minutes, after which it slowly increased. The minimum frequency is

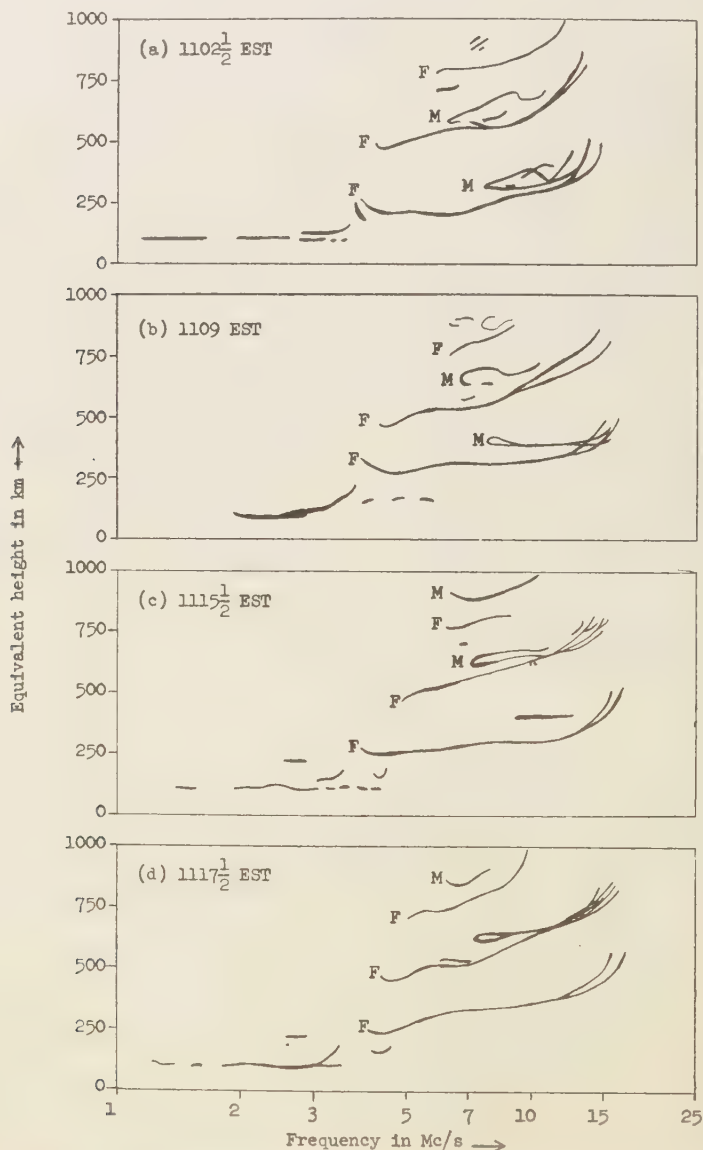


Fig. 2. Illustrating ionograms taken on Grand Bahama Island after the firing of Vanguard II. (Reproduced by courtesy of U. S. Army Signal Radio Propagation agency.)

lower for the first multiple than for the primary echo, and lower still for the second multiple.

Interpretation of observations. To interpret the unusual observations appearing in Figure 2, it is convenient to examine qualitatively the types of ionograms to be expected under various hypotheses. In doing so, we shall use the ray theory, and shall disregard magneto-ionic double refraction. We shall also disregard, temporarily, multiple echoes involving successive reflections between the earth and the ionosphere.

An ionospheric sounder is not sufficiently sensitive to obtain echoes from the skin of the missile. Nevertheless, it is convenient to begin by examining how such echoes would appear on an ionogram for a target traveling outward through the *F* region from a point in the vicinity of the sounder. As the target ascends, the series of ionograms shown in Figure 3 is obtained. In Figure 3a the target is below the ionosphere, and the horizontal broken line indicates the regular radar echo. However, the target can also be seen by reflection in the *F* region, and this echo exhibits the usual skip-distance behavior; the target can be seen by reflection in the *F* region only for frequencies less than a maximum frequency that exceeds somewhat the penetration-frequency of the *F* region. In Figure 3b it is supposed that the target is in the *F* region below the level of maximum ionization density. The ionization density at the level of the target corresponds to a certain plasma-frequency. For lower radio frequencies the target is in the over-dense portion of the ionosphere and cannot be detected by the sounder. The plasma-frequency at target-level appears on the ionogram, therefore, as a minimum frequency below which the target cannot be detected. Thus the separate target-traces in Figure 3a join together as shown in Figure 3b to form a single continuous trace. As the target ascends in the *F* region, the minimum frequency for detection increases. As the target approaches the level of maximum ionization density, the double bend in the target-trace smooths out, and when the target is above the level of maximum ionization density, the trace is as shown in Figure 3c.

Let us now suppose that the target is an ionized cloud associated with the missile. Such a cloud has a penetration-frequency that can be assumed to be greater or less than the penetration-frequency of the *F* region. The target-traces

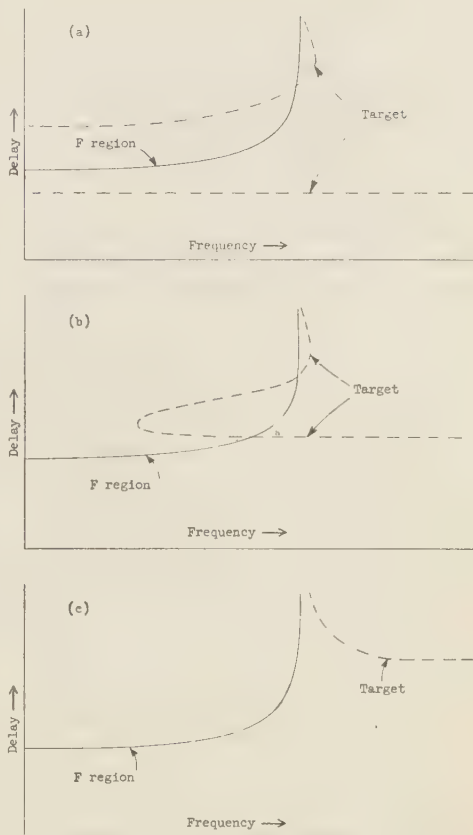


Fig. 3. Illustrating theoretical ionograms for a solid target (a) below the ionosphere; (b) in the lower part of the *F* region; (c) above the level of maximum ionization density.

shown in Figure 3 must now be modified to indicate the penetration-frequency of the cloud. In this way a series of ionograms may be sketched by assigning different penetration-frequencies to the cloud. It is clear that, in this way, we do not obtain ionograms of the type required to explain Figure 2. It is concluded that the ionograms in Figure 2 cannot be explained in terms of a persistent local increase in ionization-density at any level in the ionosphere.

In view of the heat associated with the exhaust from the missile, one would expect striking irregularities to be produced in the preexisting ionization of the *F* region. In principle, such irregularities can cause scattering, even though no increase in average ionization density may

take place. Echoes involving scattering cannot, however, be used to explain the persistent phenomenon illustrated in Figure 2. The unusual echoes in Figure 2 have traces almost as clean as regular F -region echoes. The phenomenon clearly involves total reflection, not scattering from a portion of the ionosphere.

Since the phenomenon illustrated in Figure 2 cannot be explained in terms of a local increase in ionization density, and cannot be explained in terms of irregularities produced in the normal F -region ionization, it remains to consider the possibility of a local decrease in ionization density. Let us suppose, therefore, that the hot gases associated with the missile's exhaust punch a hole in the F region, leading to contours of constant ionization density such as those sketched in Figure 4. Let us further suppose that an ionospheric sounder is located on the surface of the earth at a point not vertically below the hole. We now obtain immediately the type of ionogram that we are seeking, as shown in Figure 5a. At a sufficiently low frequency the ionization contour corresponding to the lower boundary of the over-dense portion of the ionosphere could be the lowest contour indicated in Figure 4. The tilting of this contour is insufficient to provide a 'normal' on to the ionosphere from an ionospheric sounder in the direction indicated. Thus, at sufficiently low frequencies, there is no echo associated with the hole. As we move to higher frequencies, however, the lower boundary of the over-dense portion of the ionosphere corresponds to successively higher contours in Figure 4. With a sufficiently pronounced hole, increase in fre-

quency leads to a situation in which a 'normal' from the ionospheric sounder onto the far side of the hole becomes possible. With further increase in frequency, this ray splits into two rays, giving a double trace as shown in Figure 5a. The appearance of such double traces is a standard feature of ray-tracing in a medium with a smooth distribution of refractive index; examples have already occurred in connection with the maximum and minimum frequencies for the target echoes in Figure 3a, b. With further increase in frequency, the two rays from the ionospheric sounder to the far side of the hole penetrate the F region at a frequency determined by the maximum ionization density in the ionosphere on the far side of the hole. The effect of the hole, if sufficiently pronounced, is therefore to produce an ionogram of the type shown in Figure 5a. It is immediately clear that this is the type of ionogram with which we are concerned in Figure 2.

To create an ionogram of the type shown in Figure 5a, a substantial reduction of ionization density in the hole is required, possibly of the order of a factor of two. To estimate the ionization density at the saddle-point in Figure 4 we would require an ionospheric sounder located more or less vertically below the hole. As the sounder moves toward a point roughly below the hole, the ionograms change in the manner illustrated in Figure 5b and c. For a sounder underneath the hole, one of the penetration-frequencies is appreciably less than the other two, and from this an estimate could be made of the ionization density at the saddle-point.

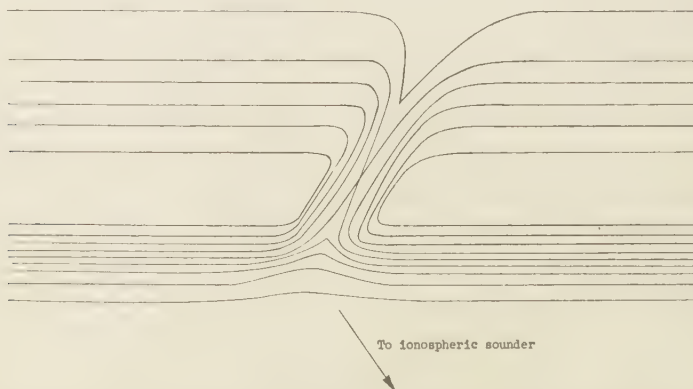


Fig. 4. Illustrating contours of constant ionization density associated with a hole in the F region.

For an ionospheric sounder located as indicated in Figure 4, let us now consider the effect of multiple reflections between the earth and the ionosphere. Multiple reflections permit a ray from the sounder to attack the hole in a direction nearer to the vertical. As we proceed to successively higher multiples, therefore, echoes from the hole should appear at successively lower frequencies. This is what happens in Figure 2. Moreover, as the hole fills up, the primary echo from the hole should disappear before the first multiple. This again is what happens in Figure 2. In the same way, the first multiple should disappear before the second multiple. This is also what happened; it is not illustrated in Figure 2 because, at this stage, echoes from the hole were weak and disappearing. If high-order multiples of echoes from the hole could be satisfactorily observed, the rays would attack the hole nearly vertically and one of the rays would have a penetration-frequency indicative of the ionization density at the saddle-point. Unfortunately such high-order multiples were not sufficiently well developed on the ionograms available.

It is concluded that the logical explanation of the ionograms shown in Figure 2 is in the terms of a local diminution in the ionization density, forming a hole through the F region qualitatively of the type sketched in Figure 4.

Discussion. It is appropriate to ask whether it is physically likely that a hole in the F region of the type required to explain the ionograms would be caused by the missile, and whether it would last for half an hour or more. The hot gases from the motor would immediately rise. Although they are ionized, recombination would be rapid. It is clear from the ionograms that ionization densities in excess of the regular maximum ionization density in the F region were never produced, except perhaps instantaneously. The rise of the exhaust gases would lead to replacement by comparatively less ionized air from below, thereby forming a hole in the F region. The hole would immediately begin to fill up by diffusion. However, at the level of maximum ionization density in the F region, diffusion of electrons across the lines of flux of the earth's magnetic field is very slow. Thus the higher ionization densities existing around the hole would produce little tendency to fill the hole by diffusion. On the other hand, diffusion along the lines of flux of the earth's magnetic

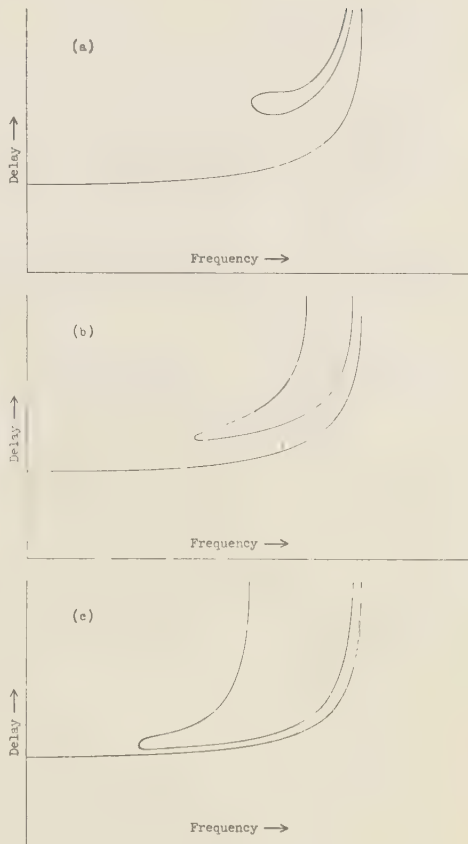


Fig. 5. Illustrating theoretical ionograms for a hole in the F region at a horizontal distance from the sounder that is (a) large; (b) small; (c) negligible.

field is far more rapid. However, this process stops as soon as the highest electron density in the tube of flux defined by the hole is the electron density at the saddle-point in Figure 4. Thus, there is a limit beyond which diffusion cannot fill the hole quickly. Since Vanguard II was fired during daylight, the principal process tending to fill the hole during the period covered by Figure 2 was probably reionization by solar radiation. The time taken for the hole to fill was of the same order of magnitude as the time taken by the F region to recover from a solar eclipse.

It would appear, therefore, that a hole was punched through the F region by the hot gases associated with the missile's exhaust, after which diffusion along the earth's magnetic field tended

to realign the hole along the lines of flux. The hole was then filled in, over a period of half an hour or more, principally by reionization under the influence of solar radiation.

A possible relation to spread F and radio star scintillation. The theoretical ionograms presented in Figure 5 in connection with the ionospheric hole sketched in Figure 4 suggest a possible explanation of the natural phenomenon of spread *F* [Wright and Knecht, 1957; Reber, 1956]. Let us consider an ionosphere in which there exists, side by side, a considerable number of holes of the type indicated in Figure 4. For a particular ionospheric sounder, some of these holes are roughly overhead and lead to ionograms of the type shown in Figure 5c. Some holes are not quite overhead and lead to ionograms of the type shown in Figure 5b. Some holes are at a considerable horizontal distance from the sounder and lead to ionograms of the type shown in Figure 5a. The aggregate of holes would therefore lead to an ionogram obtained by appropriately compounding the ionograms in Figure 5. This is by no means an unrealistic description of a spread-*F* ionogram. Such ionograms often

reveal a number of identifiable traces with different penetration frequencies, sometimes spread over a frequency-range as much as two to one. The implication is that, under spread-*F* conditions, there are numerous holes in the *F* region of the type indicated in Figure 4 with saddle-point ionization densities in the holes substantially less than the maximum ionization density in the region, sometimes by as much as a factor of four.

While identifiable traces often exist on spread-*F* ionograms, it is also common for these traces to be blurred together into a single smear with roughly the same over-all shape. To produce this phenomenon with the structure suggested above, it would be necessary partly to destroy the applicability of the ray theory. This can be achieved by reducing the horizontal dimensions of the holes, and of the spaces between holes, to a value comparable with the size of a Fresnel zone. For the frequencies concerned, the size of a Fresnel zone at an *F*-region level is of the order of a kilometer or two. A spread-*F* ionogram would therefore be produced by a distribution of ionization density of the type sketched in Figure 6.

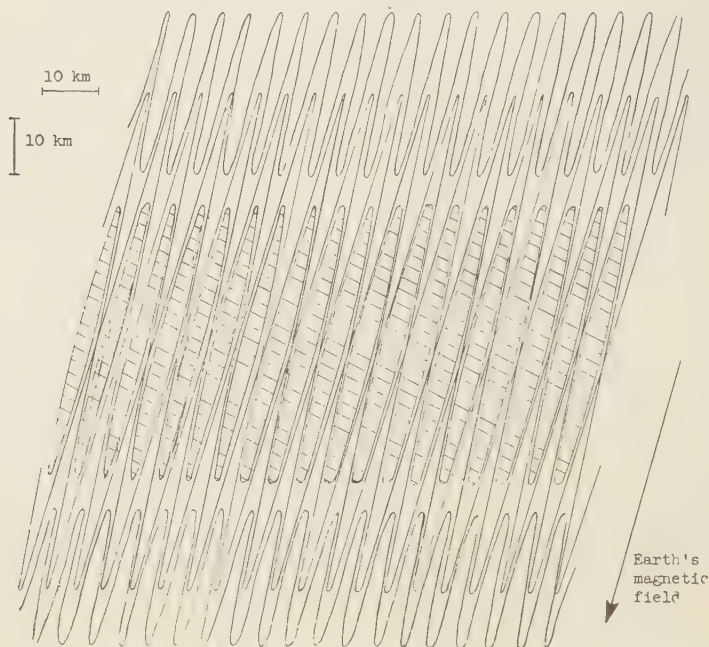


Fig. 6 Suggested structure of the *F* region for reproducing the phenomena of spread *F* and radio star scintillation.

The shaded regions in Figure 6 indicate places where the ionization density exceeds, say, $\frac{3}{4}$ of the maximum ionization density in the layer. The extreme contours at the top and bottom indicate places where the ionization density is, say, $\frac{1}{4}$ of the maximum ionization density. The holes have been shown aligned along the earth's magnetic field so that diffusion would not easily destroy them. Presumably the structure indicated in Figure 6 in a north-south direction also exists in an east-west direction.

A close relation is observed between the occurrence of spread- F ionograms and the occurrence of radio star scintillation [Dagg, 1957]. On the picture suggested in Figure 6, radio star scintillation would be due to a single layer of rather intense irregularities in ionization density aligned along the earth's magnetic field and located at the level of maximum ionization density in the F region. The scale of the irregularities measured in the direction of the earth's magnetic field would be controlled largely by the scale-height of the F region; an irregularity would probably extend further above the level of maximum ionization density than below. The scale of the horizontal structure in Figure 6 would have to fit the scale measured by Ryle and Hewish [1950], by Hewish [1952], and by Spencer [1955] in their investigations of radio star scintillation. This involves a horizontal repetition-distance of the order of 5 km as indicated in Figure 6. A distribution of ionization density in the F region of the type suggested in Figure 6 is therefore capable

of explaining both the phenomenon of spread F and the phenomenon of radio star scintillation. The fact that these two phenomena are inhibited by the presence of solar radiation [Reber, 1955; Dagg, 1957; and Booker, 1958] may also be related to the process suggested in the previous section for filling in an artificial hole in the F region produced by passage of a missile.

Acknowledgment. This work was performed by the author as a consultant to the RAND Corporation, Santa Monica, California.

REFERENCES

- Booker, H. G., The use of radio stars to study irregular refraction of radio waves in the ionosphere, *Proc. IRE*, **46**, 298-314, 1958.
- M. Dagg, Diurnal variations of radio star scintillations, spread F , and geomagnetic activity, *J. Atmospheric and Terrest. Phys.*, **10**, 204-214, 1957.
- Hewish, A., The diffraction of galactic radio waves as a method of investigating the irregular structure of the ionosphere, *Proc. Roy. Soc.*, **214**, 494-514, 1952.
- G. Reber, Worldwide Spread F , *J. Geophysical Research*, **61**, 157-164, 1956.
- Ryle, M., and A. Hewish, The effects of the terrestrial ionosphere on the radio waves from discrete sources in the galaxy, *Monthly Notices Roy. Astron. Soc.*, **110**, 384-394, 1950.
- Spencer, M., The shape of irregularities in the upper atmosphere, *Proc. Phys. Soc.*, **68**, B, 493-503, 1955.
- Wright, J. W., and R. W. Knecht, *Atlas of Ionograms Compiled for the URSI/AGI Special Committee on World Wide Ionospheric Soundings*, Boulder, National Bureau of Standards, 1957.

(Manuscript received January 30, 1961.)

On the Nature of Equatorial Spread F

ROBERT COHEN AND KENNETH L. BOWLES

*Central Radio Propagation Laboratory, National Bureau of Standards
Boulder, Colorado*

Abstract. Ionospheric propagation via scattering from the F region was sought at 50 Mc/s over a transequatorial path (with midpoint near Huancayo, Peru) employing a 2580-km transmitter-receiver separation. Propagation via F scatter was present over this path about 10 per cent of the time, though only at night. A condition closely related to the occurrence of F scatter was the presence of equatorial spread- F configurations on the Huancayo ionograms. The height of the propagation medium supporting this F scatter was usually identifiable with the lowest height of the associated equatorial spread F on the ionograms. On the basis of an interpretation of experimental results, equatorial spread F is shown to arise from scattering by relatively thin sheets of irregularities in electron density which occur at the bottom of the F layer or as much as 100 km lower. The thickness of these patches is estimated at the order of 50 km. The scattering irregularities comprising these patches are found to be elongated along the earth's magnetic lines of force. Electron irregularities of scale 10 meters or smaller measured in at least one dimension transverse to the magnetic field lines, and 1000 meters or longer measured along the magnetic field lines, are shown to exist as high as 450 km in the nocturnal equatorial F region during equatorial spread- F conditions. Equatorial spread- F echoes observed at Huancayo are demonstrated to be arriving at all elevation angles in the magnetic east-west plane. The geographical extension of a given scattering sheet in the magnetic east-west direction was at times as great as 1000 km. An observational procedure is suggested for distinguishing two fundamental varieties of spread- F echoes appearing on the equatorial ionograms. A necessary condition, that the contours of mean electron density be parallel to the magnetic lines of force, is proposed as a controlling factor for the occurrence of spread F in the equatorial ionosphere.

BACKGROUND

The basic literature in the field of spread F has not until recently been very extensive [Glover, 1960], and the experimental knowledge about this phenomenon is scanty, being largely that obtained from the study of standard ionograms [Reber, 1956; Wright and Skinner, 1959; Bowman, 1960] and from somewhat inconclusive observations of radio-star scintillations [Wright, Koster, and Skinner, 1956; Briggs, 1958; Dagg, 1957]. Much of what has been written on the subject pertains to the manifestations of spread F in the equatorial ionosphere [Wells, 1954; Booker and Wells, 1938], but in these references a distinction is not usually drawn between the several varieties of spread F encountered near the magnetic equator.

There was an opportunity to conduct studies of spread- F phenomena in the equatorial F region when, in conjunction with IGY activities of the U. S. National Bureau of Standards, a transequatorial ionospheric scatter program was initiated [Bowles and Cohen, 1957]. (The terms

'equator' and 'equatorial' are understood in this paper to refer to the vicinity of the magnetic equator.)

The feasibility of detecting propagation due to scattering in the F region appeared quite promising in the vicinity of the magnetic equator, in view of the favorable orientation there of the presumably field aligned electron irregularities with respect to any oblique propagation path centered at the equator. Also, the known presence of considerable spread F near the magnetic equator led to the anticipation of supporting ionospheric propagation by means of scattering effects related to that phenomenon.

Interest had also been stimulated in the possible importance of ionospheric scattering in the equatorial F region by the experience [Ferrell, 1951] of radio amateurs who had been able to communicate at 50 Mc/s between distant points on either side of the magnetic equator. These radio contacts had been mainly limited, however, to nighttime periods during the peak of the sunspot cycle and to equinoctial months.

DESIGN OF THE IGY EXPERIMENT

At the time of instrumenting for the IGY experiment under discussion, the spread in ranges of the equatorial spread F on ionograms was ascribed [Booker, 1956] to a mechanism whereby the time delays were introduced by a scattering screen between the F region and the ground. This hypothetical screen was estimated to be at a height of about 180 km but was supposed not to be directly detectable with an ionospheric

sounder observing the associated equatorial spread F . Also, the theory established a height 'ceiling' of 300 km for the existence of ionospheric turbulence. This turbulence was believed to be responsible for the presence of scattering irregularities causing the spread echoes. Accordingly, an F scatter experiment was designed with the antenna beam intersections arranged for an overlap in the vicinity of 180 km in height above the magnetic equator near Huancayo, Peru.

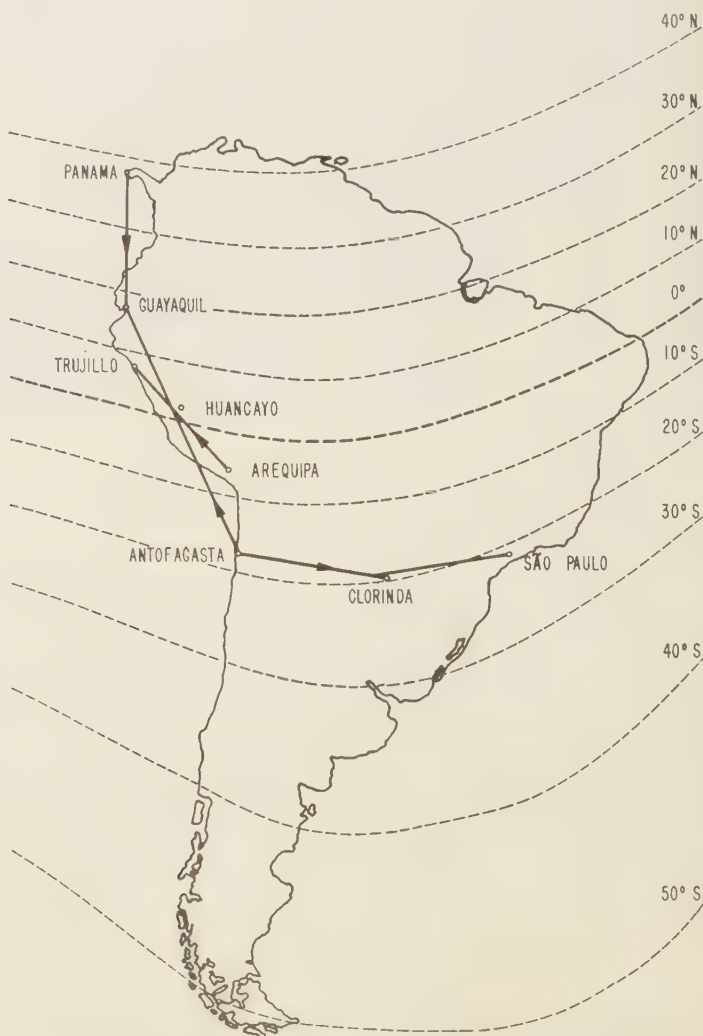


Fig. 1. The deployment of IGY forward-scatter stations in South America relative to the magnetic isoclines.

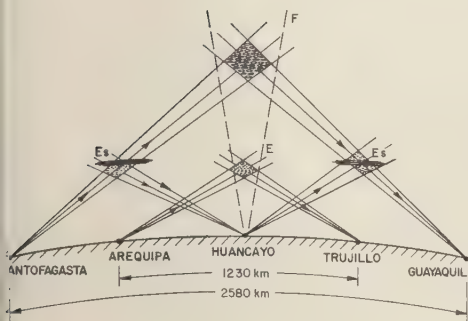


Fig. 2. Schematic diagram of the vertical cross sections of antenna beam intersections along the west coast of South America, indicating the four propagation paths referred to in this paper.

The physical deployment of all stations is shown with reference to the magnetic isoclines in Figure 1; a schematic diagram of various antenna beam intersections in the vertical plane constitutes Figure 2 (this diagram does not preserve angles).

The basic approach used in this experiment was to transmit a steady 'continuous-wave' signal and to monitor the variation in amplitude of the signal received at a certain distance from the transmitter. The F scatter experiment was designed so that the received signal would usually arrive via this mode of propagation. Thus, the temporal variations of the received signal could be related to changes in the propagating medium. By means of such techniques as transmitting pulses of radio energy over this path, information was obtained about the height of the propagating medium.

All the radio circuits in these experiments were operated within a few kilocycles per second of 50 Mc/s. In order to study F scatter while excluding single-hop scatter signals propagated via the E region, a separation of 2580 km was employed between a transmitter located at Antofagasta, Chile ($23^{\circ}44'S$, $70^{\circ}15'W$), and a receiver located near Guayaquil, Ecuador ($2^{\circ}36'S$, $80^{\circ}24'W$). This separation gave a horizon cutoff height, at the path midpoint, of 120 km. In view of the low signal strengths of the scatter propagation produced by irregularities of electron density in the ionosphere, it was necessary to make use of large, directive antennas (normally, rhombics 50 wavelengths long) and powerful transmitters (2 to 20 kw) in this investigation.

GEOGRAPHIC AND ANTENNA BEAM CONSIDERATIONS

The initial antenna beam overlap volume of the rhombic antennas employed is delineated in Figures 3 and 4 by various cross-sectional cuts. Figure 3 is a diagram of cross sections obtained by passing horizontal surfaces through the scattering volume at selected heights and computing the loci at the levels of points in the product antenna pattern that are 20 db weaker than the maximum sensitivity attainable. The left-hand part of this diagram indicates the horizontal extent of the scattering volumes for three basic E -region paths; the right-hand part shows the horizontal extent of the F -region scattering volume, based upon directing the antennas at the height neighborhood from which F scatter was originally anticipated.

The midpoint of the oblique F scatter path is about 110 km south of Huancayo ($12^{\circ}3'S$, $75^{\circ}20'W$) and almost exactly at the magnetic equator. Therefore, scatter propagation over this path via field aligned irregularities would be expected to occur principally in the ionosphere above the magnetic equator. (This conclusion results from the requirement of a 'mirroring' geometry of the transmitter and receiver locations relative to the magnetic field aligned irregularities, in order to have phase coherence of the scatter signal contributions from individual irregularities.) On the other hand, from orthogonality considerations, the field aligned F -region irregularities detectable at Huancayo will be those in an east-west plane passing through the F region 10 to 15 km north of Huancayo. The two systems of observations are therefore sensitive to irregularities in somewhat different geographical locations. Thus, inasmuch as the spread- F irregularities may occur in patches of limited horizontal extent, complete correlation between F scatter effects observed with the Huancayo ionosonde and at oblique incidence in Guayaquil should not be expected. To obtain some measure of this extent, spread- F features on the Huancayo ionograms can be compared with those obtained concurrently, using identical equipment, at Chimbote ($9^{\circ}4'S$, $78^{\circ}35'W$), located about 450 km northwest of Huancayo. The differences in these features are found to be appreciable, so that the effect of the 125-km separation of the VHF scattering volume from



Fig. 3. The west coastal area of South America with the approximate loci of antenna beam intersections computed for the height (in kilometers) adjacent to each locus. The left-hand diagram is for the three propagation paths in the *E* region; the right-hand diagram is for the long path (before raising the antenna beams).

the ionosonde scattering volume may at times be significant [Knecht, 1960]. (Some techniques for studying the east-west extent, thickness, and drift velocity of these scattering patches in the equatorial *F* region are discussed in a sequel [Calvert and Cohen, 1961] to this paper.)

Figure 4 is a representation along the vertical plane passing through the great circle between Antofagasta and Guayaquil. The beam intersections to 6-db points in the joint antenna patterns of the main lobes are shown by the diagonal shading, along with the possible regions (finely dotted) in which propagation might be sustained through the combination of the main lobe of one antenna with the second lobe of the other. The auxiliary scattering regions so defined are to the north and south of the path midpoint and may be contributing from time to time to the scatter signal if there are appropriately

oriented irregularities of sufficient strength.

During the final weeks of the year of observation, the antenna beams were swung upward by lowering the antennas to half-mast. The main lobe intersections for the rhombic antennas after lowering them are depicted in the coarsely dotted part of Figure 4.

The scattering region to which the system was most sensitive before the change was below about 300 km for the principal path, with weak effects possible, through the participation of the second lobes, up to about 320 km. After the antenna change, it was possible to observe *K* region scattering above the path midpoint up to about 450 km.

(It should be explained that the main lobe width in the vertical plane before raising the beams was constricted to about 3°, owing to the effect of the image rhombic at this lo-

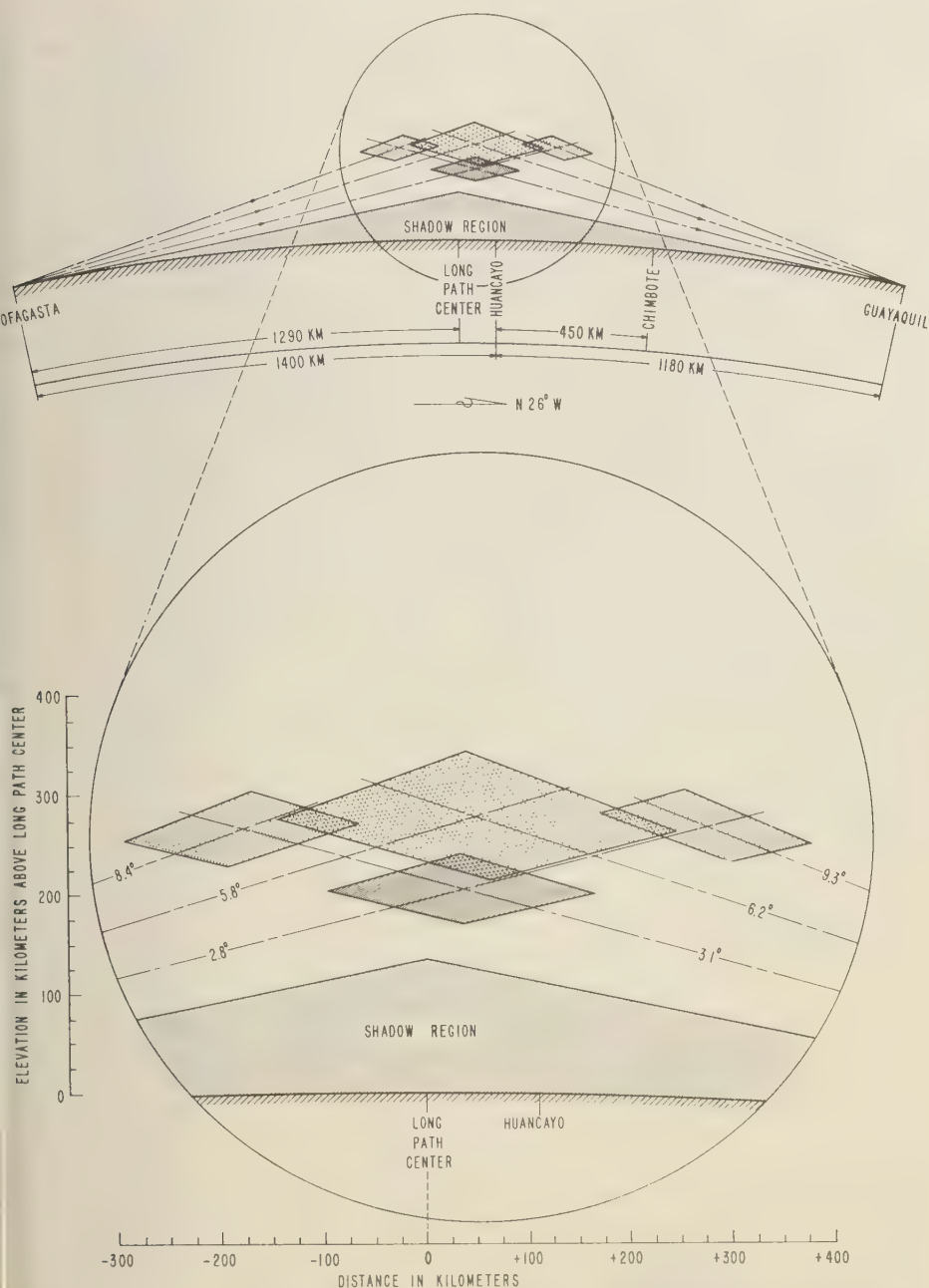


Fig. 4. Detailed antenna beam intersections in the vertical plane along the Antofagasta-Guayaquil great circle, before and after lowering the rhombic antennas.

takeoff angle, whereas the free-space rhombic beams are normally 6° wide in both the vertical and the horizontal planes.)

RESULTS OF THE CONTINUOUS-WAVE MEASUREMENTS

General discussion. In spite of the horizon-cut-off's being placed at 120 km in an effort to suppress single-hop *E* scatter over the Antofagasta \rightarrow Guayaquil path, the scatter effects in the equatorial *E* region proved to be so intense that an appreciable signal was received at Guayaquil via this mode for much of the time. It was strongest during daylight hours, and corresponded at these times (except during the presence of temperate-latitude sporadic *E*) in almost complete detail with the signal propagated over the *E* scatter path from Arequipa, Peru to Trujillo, Peru (cf. Figs. 1, 2, and 3). From time to time, after sunset in the *F* region, evening enhancements were noted in the Antofagasta \rightarrow Guayaquil records that were not apparent in the Arequipa \rightarrow Trujillo records (nor were there concurrent indications of two-hop sporadic-*E* propagation via Antofagasta \rightarrow Huancayo thence Huancayo \rightarrow Guayaquil, both of which paths were also under surveillance). Inasmuch as the Arequipa \rightarrow Trujillo path involved an antenna-pattern intersection (cf. Fig. 3) with a ceiling of about 120 km whereas the horizon planes impose this height as a lower limit for the Antofagasta \rightarrow Guayaquil scatter propagation, we interpret these

observations to mean that scattering during these openings must have been sustained at altitudes higher than 120 km. The characteristic early evening enhancements bear an almost 1:1 correlation with the appearance on Huancayo ionograms of equatorial spread *F*. A typical occurrence of one of these enhancements (cf. Fig. 5) was at 2010 hours on January 19, 1958. The Huancayo ionograms just before 2010 hours and shortly after that time differed considerably, in that equatorial spread *F* was absent or present, respectively.

In summary, for one calendar year of recording over the Antofagasta \rightarrow Guayaquil path, the only manifestations of *F* scatter seem to have occurred at night, more commonly before midnight. The preponderant mode of propagation over this circuit is by one-hop *E* scatter, via equatorial sporadic *E* in the daytime. One- or two-hop *E* scatter propagated via temperate-latitude sporadic-*E* formations appropriately situated occasionally occurs on this path.

Identification of F scatter. Some characteristics of the propagation over the 'long path' (Antofagasta \rightarrow Guayaquil) may be studied by comparing monthly hourly median signal strengths on that circuit with the corresponding plots for the 'short path' (Arequipa \rightarrow Trujillo). This comparison is shown in Figure 6, where the long-path curves have for convenience been shifted upward by 30 db, which is the usual daytime signal-strength difference between the two paths. (The reference level, 0 db, is to 1 microvolt open-circuit voltage appearing at the terminals of a 50-ohm impedance unbalanced transmission line matched to the receiving antenna. Except during special experiments antennas for both transmitting and receiving were conventional rhombics, 50 wavelength long, of a design employed earlier [Bailey, Bateman, and Kirby, 1955]. All signal strength analyzed have been normalized to levels that would have been received had the transmitter power been 2.0 kilowatts.) We note in Figure 6 the qualitative daytime correspondence between the two paths due to scattering in the *E* region and the evening rise, associated principally with *F* scatter, of the long-path signal relative to the short path. (There has been no attempt to exclude from the long-path statistics signals arising by one or two hops from temperate-latitude sporadic *E*, which may contribute some

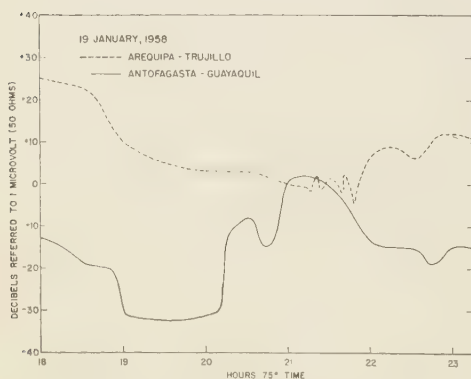


Fig. 5. Simultaneous signal strengths for the long and short paths on January 19, 1958, indicating a signal enhancement on the long path associated with equatorial spread *F* at Huancayo. (The decibel levels are not absolute.)

eat to this systematic evening enhancement.) A search has been made on a daily basis for evidence of possible F scatter over the long path during daylight hours, when there might occur regularities in the F region which could not be seen with the ionospheric sounder. At times of unusually weak propagation over the long path, the equatorial sporadic- E mode, which would otherwise be expected to obscure any daytime scatter signal over the long path, a comparison of the long-path records with those for the short path indicated that there were no residual daytime signals on the long path attributable to F scatter.

Oblique scatter observations compared with ionograms. During the year in which oblique scatter observations were conducted over the long path, the nighttime F scatter was apparent some 10 per cent of the total transmission time. Although the occurrence of this F scatter was closely correlated with the equatorial spread F (range type spread), there was no similar correlation with the unambiguous presence of temperate-latitude spread F (frequency type spread) on the Huancaño ionograms.

To exhibit the detailed correspondence between oblique scatter propagation and the various spread- F configurations appearing on the Huan-

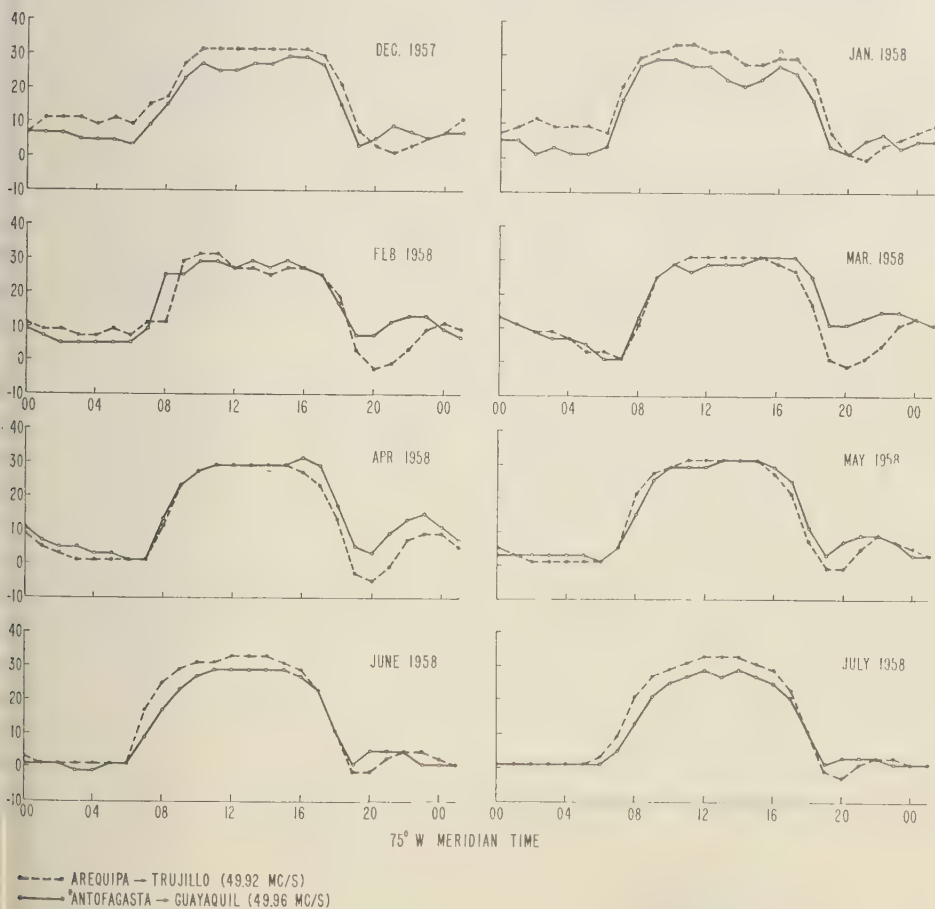


Fig. 6. A monthly comparison of the diurnal variation of hourly median signal strengths over the long and short paths, with the long-path signal shifted +30 db.

cayo ionograms, bar graphs have been prepared for the 1800 to 0600 hours period throughout the data year (Figs. 7, 8, and 9). Three categories of spread F are indicated on these charts: (1) that clearly of the range spread or equatorial variety; (2) that clearly of the frequency spread or temperate-latitude variety; (3) that which resembles frequency type spread F but which has vestiges of horizontal features such as 'spurs,' or evidences a gap between the spread region and the normal F trace. (When even the slightest ambiguity existed in the classification of frequency type spread, this third category [Calvert and Cohen, 1961] was used in preference to the second.)

Whether the minimum virtual height or 'bottom' of the nocturnal F layer is above or below 450 km is also displayed in these figures for the first and third categories of spread F . This height, which is an upper limit to the lowest height of equatorial spread F , is indicative of the likelihood of oblique scatter, on the basis of geometrical considerations related to the antenna beam intersection volume (see below). The simultaneous occurrence of the first and third spread- F categories is shown in Figures 7, 8, and 9, but, owing to graphical limitations, if the first and second categories coexist in time, only the range spread variety is denoted.

Signal enhancements over the long path are indicated in Figures 7, 8, and 9 as being either definitely due to F scatter or somewhat ambiguous. Considerable subjectivity necessarily enters the determination as to when an F scatter opening occurs, especially since there are competing modes of propagation that have to be eliminated. These modes, principally single- or double-hop sporadic- E propagation, can obscure the usually weaker F scatter signals that may be present concurrently. Thus, it may not even be possible to recognize an F scatter signal on account of uncertainty in resolving it from contaminating modes of propagation. In identifying openings, reliance has been placed, to some extent, upon the shape of the time variation of the scatter signal amplitude, inasmuch as certain characteristic patterns of the spread- F evolution may be recognized.

In the final analysis of the correspondences to be anticipated and of those actually obtained, it should also be recalled that the scattering regions viewed by the ionosonde and over the oblique path are not quite identical geograph-

ically, as was discussed earlier, and that the third spread- F category is an ambiguous classification. In particular, since the minimum height (assumed to be identical with the minimum virtual height) of the nighttime F layer is often higher than 450 km above Huncayo, and inasmuch as the antenna beam intersection was effectively limited (cf. Fig. 4) during most of the year to have a 'ceiling' below that height, there was less possibility of observing oblique scatter enhancements when spread F occurred at such times.

Angle of arrival of oblique scatter signals. Recordings over the Antofagasta → Guayaquil ' F scatter' circuit began in December of 1957. In August 1958 it was discovered at Guayaquil that, during an evening enhancement of the class previously described, a signal comparable to that being received on the rhombic antenna was available from a dipole antenna much closer to the ground. This was an indication that the signals were arriving at higher angles than had been predicted at the time of designing the experiment. Efforts were thereupon made to monitor the received signals simultaneously on various antennas at several elevation angles, and a typical evening recording displaying such comparison during equatorial spread- F condition is shown in Figure 10 for September 20, 1958. We note the nearly comparable signal strength received by the several antennas, the greater detail viewed by the higher-gain antennas, and the transformation of the F propagated signal into an essentially E scatter signal after 21½ hours. In late September, shortly after the recording was taken, the rhombic antennas at Antofagasta and Guayaquil were lowered to half-mast so as to look at a higher common volume—centered at about 300 km—more in accordance with the region from which the available F scatter was apparently being propagated. This antenna arrangement was continued for the duration of the experiment, that is, until the end of November 1958, and was employed for the pulse observations taken in that month.

RESULTS OF THE PULSE MEASUREMENTS

Purpose of the pulse measurements. As a means of establishing more precisely the region in the ionosphere sustaining the propagation enhancements associated with equatorial spread F seen at Huncayo, instrumentation was installed at Antofagasta and Guayaquil during November

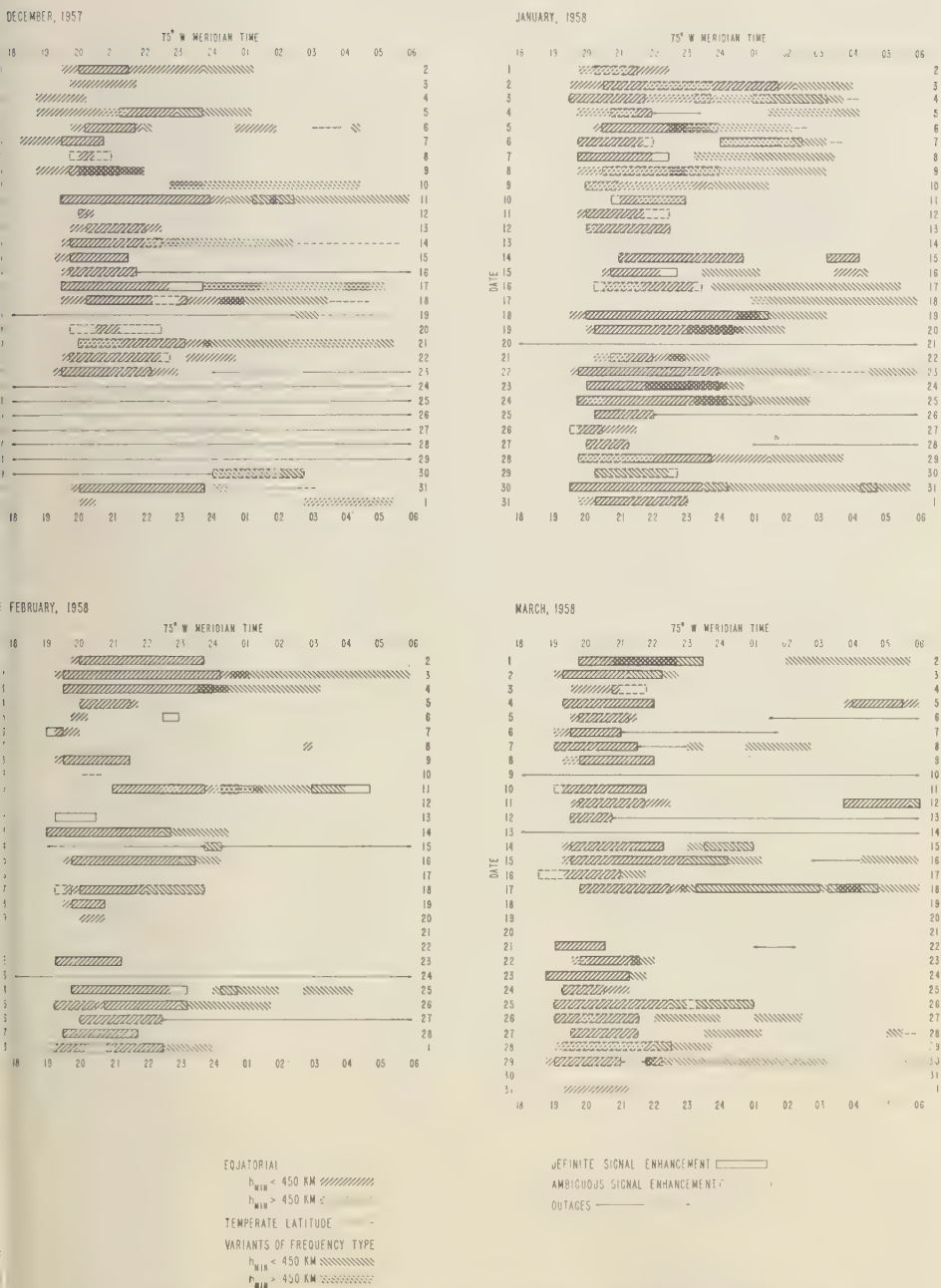


Fig. 7. The nocturnal time variation of spread-*F* configurations appearing on Huancayo ionograms and that of associated signal enhancements occurring over the Antofagasta → Guayaquil scatter path, for the first 4 months of the data year.

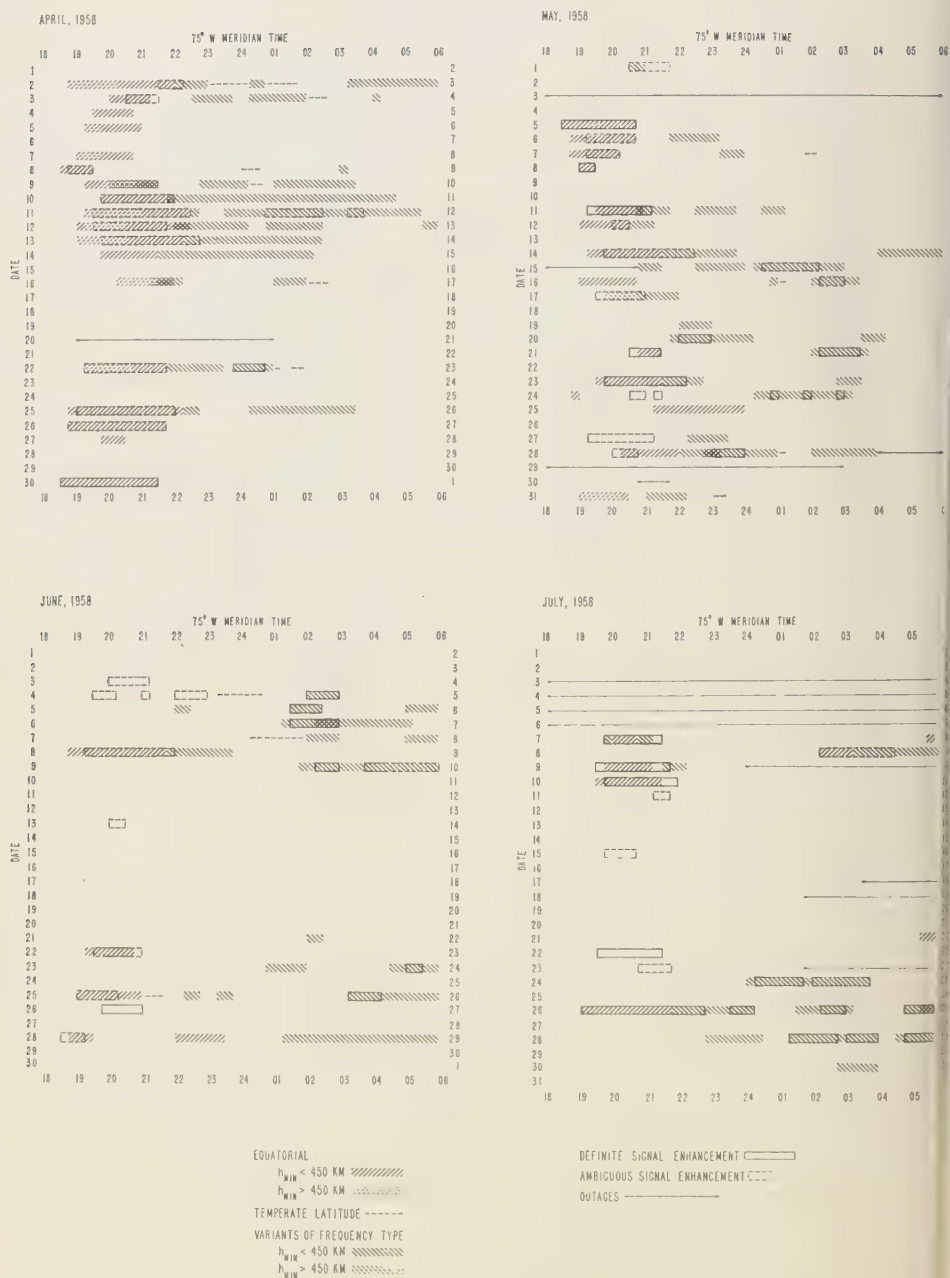
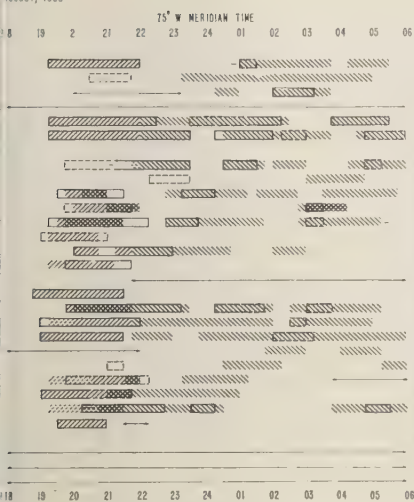
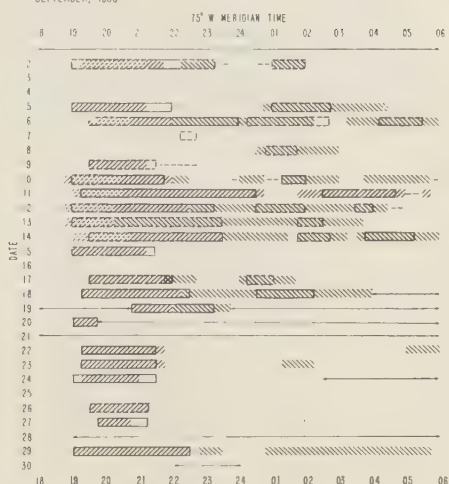


Fig. 8. The same for the second 4 months of the data year.

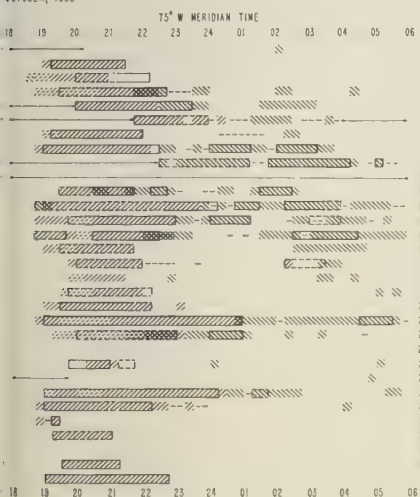
AUGUST, 1958



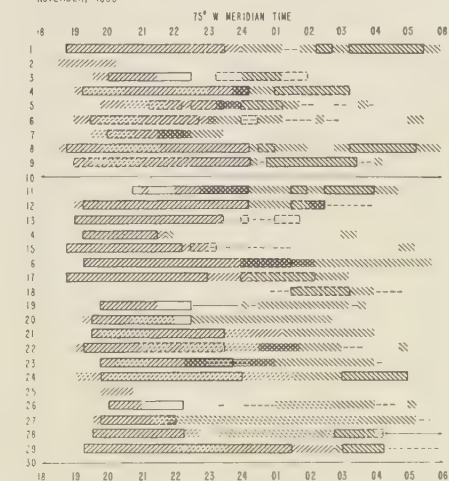
SEPTEMBER, 1958



OCTOBER, 1958



NOVEMBER, 1958



EQUATORIAL

 $f_{min} < 450 \text{ km}$ // $f_{min} > 450 \text{ km}$ //

TEMPERATE LATITUDE - - -

VARIANTS OF FREQUENCY TYPE

 $f_{min} < 450 \text{ km}$ // $f_{min} > 450 \text{ km}$ //

DEFINITE SIGNAL ENHANCEMENT

AMBIGUOUS SIGNAL ENHANCEMENT

OUTAGES

Fig. 9. The same for the third 4 months of the data year.

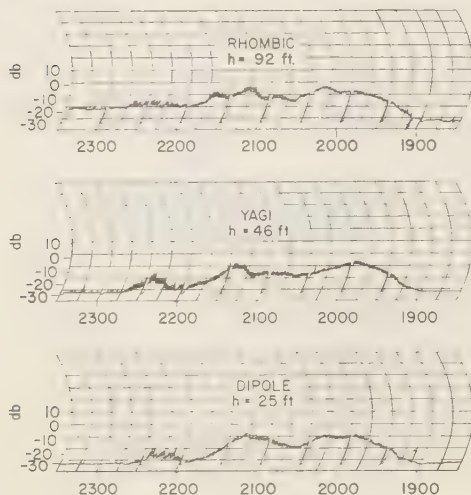


Fig. 10. Signal strengths received at Guayaquil from Antofagasta on September 20, 1958, at increasing angles of arrival, using various receiving antennas. (The decibel levels for the three antennas are not absolute but relative to an identical reference level.)

1958 for the transmission and reception of pulses. The pulse tests involved transmission in only one direction, so that the results of these measurements could give only the *relative* heights of the propagation media. Since signals propagated

from Antofagasta to Guayaquil via a single hop in the *E* region were limited by the horizon planes of transmitter and receiver to scatter from a height of at least 120 km, and on account of the limited vertical extent of the equatorial sporadic *E*, it is a reasonable assumption to make pulse delay comparisons on the basis of a reference height level of 120 km for the *E* propagated signals. A straightforward trigonometric calculation then permits the estimation of absolute heights for any ionospheric irregularities producing scatter signals having greater delays. On the basis of a one-hop or two-hop scatter mode, Figure 11 indicates the result of such a computation. The determination as to which scatter mode predominates over the Antofagasta → Guayaquil circuit at a given instant is facilitated by reference to records taken between Antofagasta and Huancayo, Huancayo and Guayaquil, and Arequipa and Trujillo.

Interpretation of the pulse delays. The computation of Figure 11 is based upon the assumption that the height delays are introduced as a result of vertical displacements of scatterers at the path midpoint. It is important to estimate the sensitivity of the pulse delay measurement to horizontal displacements of the scatterers. As was stated earlier, it would not be expected from the geometrical symmetry requirement that scattering can occur very far to the north or south of

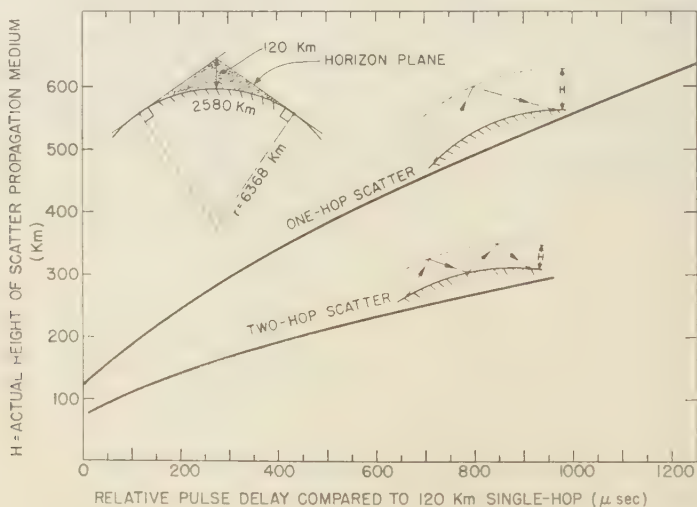


Fig. 11. Time delay of one-hop and two-hop scatter over the Antofagasta → Guayaquil path relative to the delay for single-hop scatter at a height of 120 km.

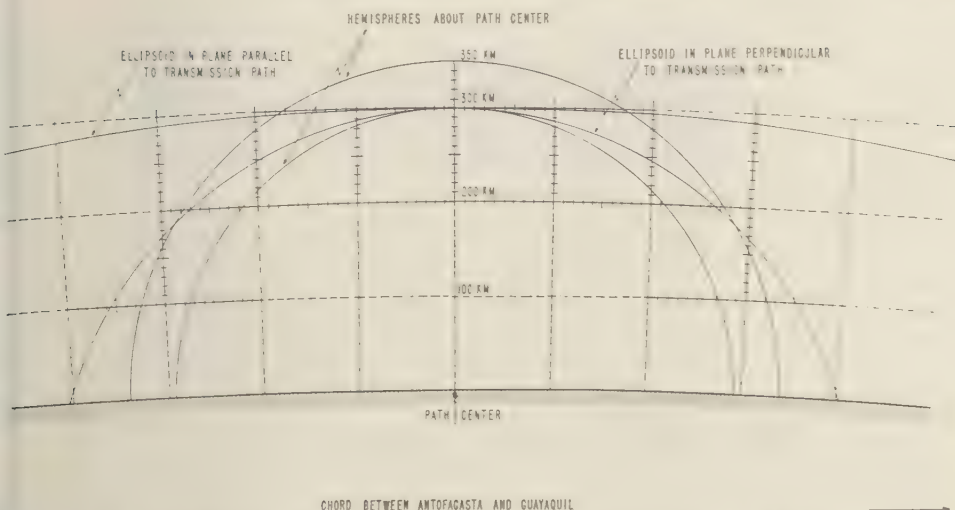


Fig. 12. A representation of a constant time delay ellipsoid in planes parallel and perpendicular to the great-circle path between Antofagasta and Guayaquil.

the path midpoint. For scatterers displaced to the east or west of this point, however, scattering is likely to occur. To study how horizontal displacements affect the observed pulse delays, an ellipsoid of revolution having the transmitter and receiver locations as foci is considered. The surface of such an ellipsoid is the spatial locus of constant time delay. Figure 12 is a representation of an ellipsoid of this kind corresponding to a height of 300 km above the path center. Cross sections of the ellipsoid are shown in the plane of the transmission path and in the plane transverse to the path. Also, hemispheres at 300 and 350 km from the path center, shown in projection, correspond to the range from a symmetrically located sounder. (Actually, the Huancayo sounder is not situated quite so ideally.) Then, upon considering displacements along the transmission path, it is apparent that they constitute a second-order effect in interpreting pulse delays. It is evident from Figure 12, however, that the oblique measurement is considerably more sensitive to transverse displacements, and the sensitivity of an ionosonde at the midpoint to horizontal displacements is seen to be even higher.

Methodology of the pulse measurements. The possibility of carrying out relative pulse delay measurements is assured, even at times when the sporadic E propagated reference signal is not

available, because of the high stability (1 part in 10^8 per day) of the frequency standards employed for pulse and sweep synchronization at the transmitter and receiver, respectively. Once a reference delay has been established for single-hop E scatter, it may be extrapolated for many hours for comparison purposes.

Simultaneous pulse and continuous-wave recordings were obtained by maintaining the normal radio-frequency excitation of the transmitter at 49.96 Mc/s while transmitting pulses having a radio frequency of 49.98 Mc/s, just 20 kc/s displaced. (The 0.04 per cent separation in these exploring frequencies is presumed to produce a negligible difference in the propagation effects.) Then, by utilization of the regular continuous-wave receiver and an auxiliary pulse-adapted receiver adjusted to the optimum bandwidth (15 to 20 kc/s) for the 60-microsecond pulses that were being transmitted, both a 'range-time' presentation on an oscilloscope and the regular 'amplitude-time' paper tape record were obtained. The inherently greater system sensitivity for the continuous-wave recordings compared with that for the pulse records is apparent in the examples shown in Figures 13, 14, 15, and 16.

Analysis of the pulse results. The Huancayo ionograms for significant times during the intervals displayed in Figures 13 to 16 are

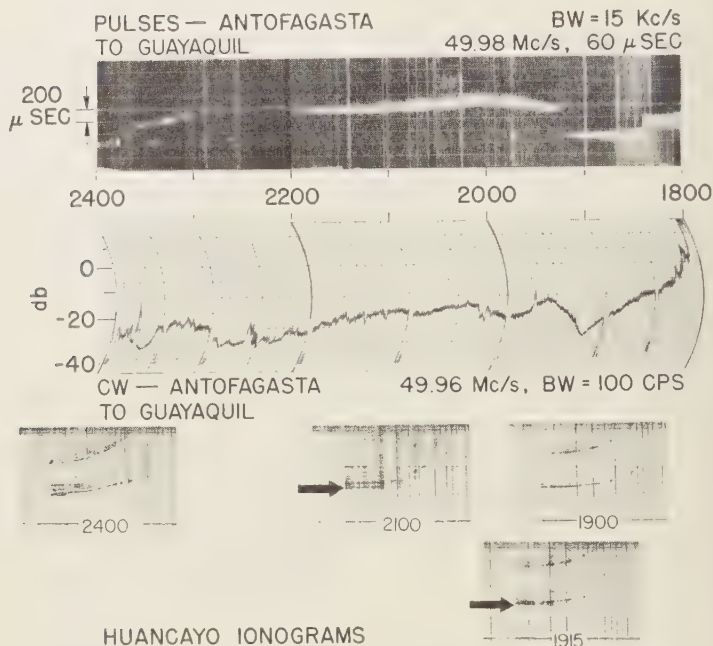


Fig. 13. Simultaneous recordings on November 16, 1958, of pulse delay and signal strength over the Antofagasta → Guayaquil path, with selected Huancayo ionograms.

positioned in accordance with the time at which each sounding was taken. When a time delay determination for a spread- F opening can be established from the pulse records by means of Figure 11, an arrow is superimposed on the associated ionogram at the height appropriate to the measured time delay.

Proceeding chronologically, let us examine some typical examples of the records obtained in this way. Figure 13, for November 16, 1958, is a representative example of the time development of the equatorial spread- F scatter intensity. The signal being received at 1800 is interpreted as being propagated by equatorial sporadic E . As E -region sunset approaches, this signal gradually drops out. (The apparent change in pulse delay of this signal at about 1822 hours resulted from an equipment adjustment made so as to place the signal in a more appropriate position on the pulse record.) At about 1910 a rather pronounced rise in signal strength is noted on the continuous-wave record. It is seen to correspond to the transition on the Huancayo ionograms from a smooth F trace at 1900 to an equatorial spread- F configuration at 1915. The time delay

measured at 1915 relative to the E propagated signal results in an estimation of about 370 km for the height of the scattering region. This coincides with the lowest edge of the equatorial spread- F return on the ionogram. (The frequency at vertical incidence equivalent (see below) to the 50-Mc/s frequency over the oblique path with scattering at 370 km comes to about 15 Mc/s. This means that, using sufficient sensitivity with the sweep-frequency radar at Huancayo, the lower edge of the equatorial spread- F echo on the ionogram would extend to at least 15 Mc/s on the frequency scale, considerably beyond the F critical frequency.) Similar results are obtained at 2100, when the scatter region has risen slightly from its height at 1915. Throughout this evening, vestiges are apparent on the pulse record of an E -region signal which maintains a relatively fixed time delay. At 2400, when the Huancayo ionogram indicates only satellite traces and mainly spread F of the temperate-latitude variety, there is little evidence of an F propagated scatter signal.

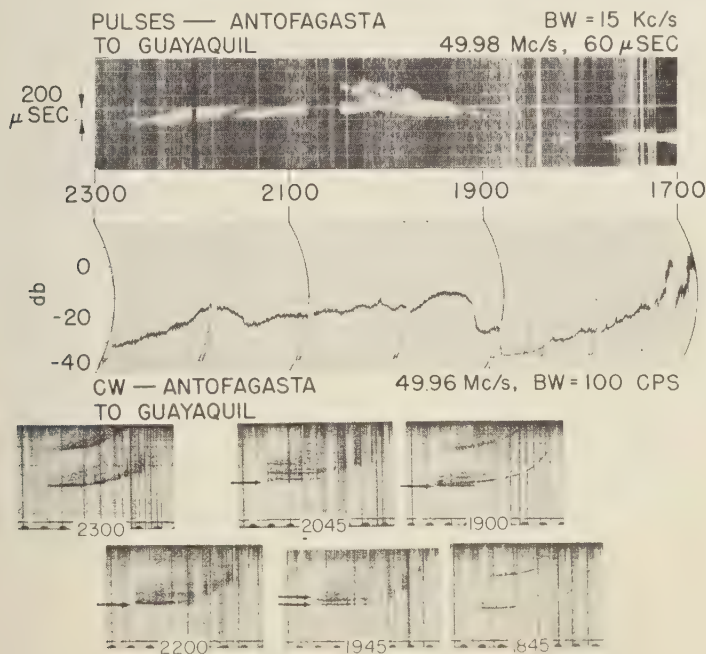
The record of Figure 14 for November 16, 1958, tells a similar story, but this time somewhat

more intricate behavior of the F scatter is in evidence. (An equipment failure caused the gap in the pulse record from 2028 to 2050.) At least two time delays are associated with the F propagation between 1930 and 2030, suggesting the presence of several strata of scattering irregularities. At 2300, when the evening enhancement was characteristically abated, the equatorial spread F has been almost completely replaced by temperate-latitude spread F , with little or no associated F scatter.

On the night of November 22, 1958 (cf. Fig. 15), there is first a bifurcation of the F scatter signal; then the lower trace gradually rises, disappears as pulse delays corresponding to scatter heights up to 450 km and above are approached, then reappears. The several traces are indicative of multiple strata of spread- F irregularities, and the disappearance of the pulse indications at 450 km is not surprising, inasmuch as the antenna beam intersection is in a null at that level. Also, higher scatter angles and longer ray paths result in a reduction of the amplitude

of scattering at such heights. Figure 15 provides another demonstration of the lack of F scatter at times (such as 2400 here) when temperate-latitude spread F predominates.

An apparently anomalous class of spread F . A class of enhancements was observed during the early morning hours, generally for an hour or two, which at first glance seems to be associated with frequency type spread on the Huancayo ionograms. On closer examination the ionograms corresponding to these openings are recognized to be somewhat more complex, frequently exhibiting a regular F trace with a separated scatter region. An example is given in Figure 16 for the early hours of November 25, 1958. The peak of activity at about 0400 corresponds to just such a complex ionogram, and the height of scattering can be approximately identified with that of a horizontal stratum apparent in what otherwise resembles temperate-latitude spread F . Below this stratum is a noticeable gap between the spread and the main trace. Also present is a spur, at about the lowest height of



HUANCAYO IONOGRAMS

Fig. 14. The same on November 17, 1958.

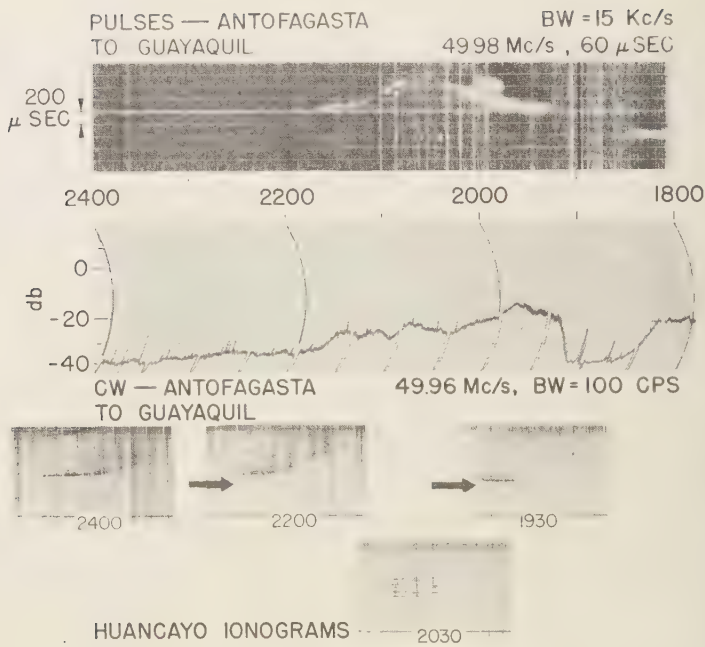


Fig. 15. The same on November 22, 1958.

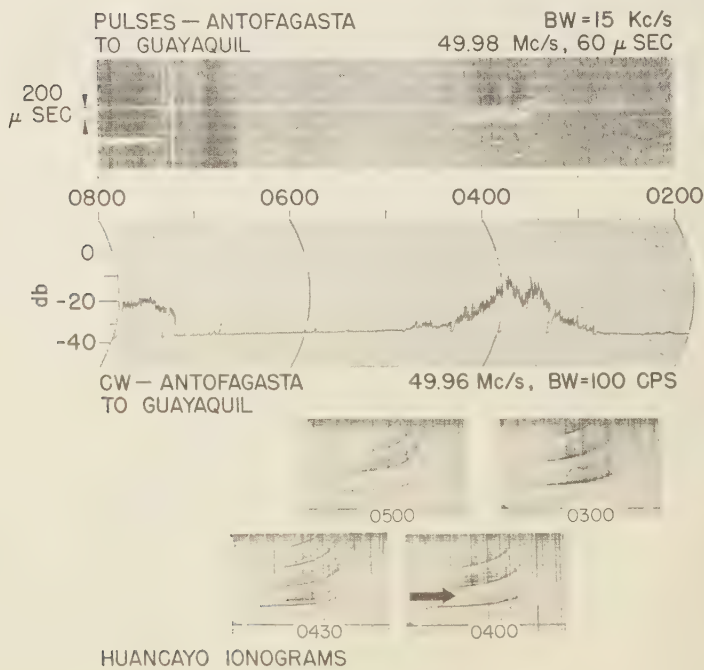


Fig. 16. The same on November 25, 1958.

the spread, extending beyond the critical frequency. This configuration can be associated with scattering irregularities embedded in the F layer and viewed to the east or west of Huancayo; this general case is treated in another publication [Calvert and Cohen, 1961]. A branching of the spread echo is frequently evidenced at its low-frequency limit, as in this example at 9400 and 9430.

Summary of the pulse results. In these examples there is: (1) generally excellent agreement of the lowest height of the equatorial spread- F configuration with the height deduced from the pulse delays, and (2) absence of F scatter propagation when only temperate-latitude spread F is present.

An additional uniformity is exhibited by all the pulse observations—the lack of appreciable broadening of the 60-microsecond VHF pulses in the course of their scatter propagation. This result, combined with the interpretation of another experimental observation to be described shortly, leads to the conclusion that the irregularities in the F region associated with equatorial spread F occur in fairly thin sheets. Also, from what has been seen of the correspondence of the lowest height of the equatorial spread- F configuration on Huancayo ionograms and the F scatter heights, it seems clear that the spread- F time delays are not introduced by an intervening forward-scattering screen considerably below the F layer but *invisible* on the ionograms, as had been suggested from theoretical considerations [Booker, 1956]. This conclusion follows from the demonstration here that the equatorial spread F arises from a scattering sheet actually at the base of the F layer, or below, and identifiable with the lowest height of the spread on the ionogram. If the intervening-screen theory were correct, the observed time delays over the oblique path could be no less than those associated with normal F reflections. From Figure 11, using a representative pulse broadening associated with a typical pulse delay, it is estimated that the thickness of this zone is of the order of 50 km. This result is corroborated in the next section.

CONCLUSIONS FROM BOTH SETS OF OBSERVATIONS

From the detection of radio wave scattering across the magnetic equator at heights approaching 450 km, it is possible, assuming a scattering mechanism, to deduce an upper limit to the scale

size of the irregularities responsible for this scattering. This deduction is based upon the expected occurrence, in a theoretical treatment of the F scatter power dependence, of the geometrical parameter $\lambda/(2 \sin \theta/2)$. By computing a representative value for this parameter, where $\lambda = 6$ meters and θ , the scatter angle, is about 35° in the present oblique-scatter case, it may be inferred that the scale size of the irregularities that must exist in the equatorial F region up to 450 km or more is 10 meters or smaller, measured transverse to the magnetic field lines. From the invariance of this parameter may be deduced the equivalent wavelength at vertical incidence, where, with $\theta = 180^\circ$, we require that $\lambda = 20$ meters, corresponding to the frequency of 15 Mc/s previously alluded to. (Radar measurements at Huancayo with a high-gain vertically directed antenna have established that equatorial spread- F echoes can indeed be detected, with sufficient sensitivity, above the F critical frequency.)

MAGNETIC FIELD ALIGNMENT OF THE SPREAD- F IRREGULARITIES

Introduction. It is expected on theoretical grounds that ionospheric irregularities above about 100 km will be aligned with the magnetic field. Experimental studies have been conducted at Huancayo on equatorial spread- F echoes in order to verify that the associated irregularities are, indeed, field aligned. In the following paragraphs two experiments confirming the theoretical expectation are described.

Antennas for establishing field alignment. For these experiments a pair of vertically beamed broadside antenna arrays was erected at Huancayo. These antennas each consisted of 6 cophased full-wave (at 10 Mc/s) dipoles arranged in a 3×2 pattern, suspended about a quarter-wavelength above the ground by a common 3×3 array of masts. One antenna, to be referred to as antenna 1, was oriented with its dipoles in the magnetic east-west direction; antenna 2 was installed orthogonally to 1. The antennas could be used over a frequency range extending 2 Mc/s above and below the 10 Mc/s design frequency. They were fed in such a way that, by means of a switching system and a duplexer, either antenna could be used, in radar operation, for both transmitting and receiving, or one antenna could transmit and the other receive. Thus, four modes

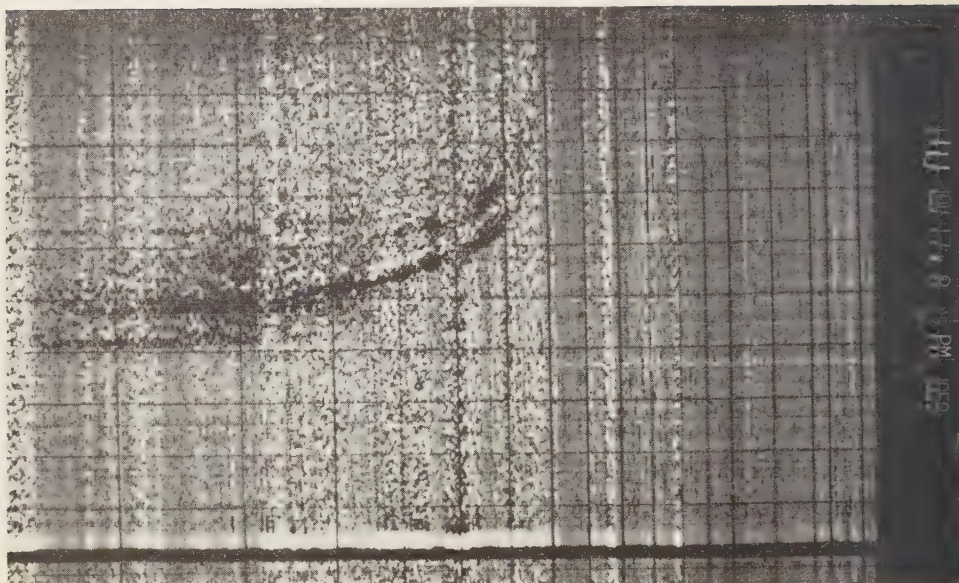


Fig. 17. The Huancayo ionogram for 2025 hours on November 17, 1959, with 100-km height markers.

in all were available, designated as 1-1, 2-2, 2-1, and 1-2, where the antenna first mentioned is the transmitting one. For these tests the antenna arrays were always used in conjunction with the C-4 ionosonde, generally at fixed frequencies in the 8 to 12 Mc/s range.

The gain of the antenna arrays described is considerably greater than that of the delta antenna normally employed with the C-4 at Huancayo. Indeed, the beamwidth at 10 Mc/s is about 25° to 30° for the broadside arrays, compared with about 60° for the delta. The beam-narrowing capabilities permit of some interesting comparisons in making spread- F studies, as will shortly be demonstrated. The broadside arrays have the additional property that they may be operated in a split-lobe mode whereby they possess a null in the vertical direction and radiate two broad lobes, each maximizing 30° from the vertical, to the east and west for the east-west polarized antenna 1, and to the north and south for the north-south polarized antenna 2. This beam-splitting feature permits a direct demonstration that the irregularities of interest are magnetic field aligned.

Results of the lobe-switching observations. First

we discuss a *direct* proof of field alignment, based upon the results of employing the beam-splitting technique. Figures 17 and 18 indicate, respectively, the pertinent observations at Huancayo utilizing the ionospheric sounder in conjunction with the delta antennas, in the normal sweep-frequency presentation, and with the broadside antenna array operated at the fixed frequency of 8.4 Mc/s. The four A-scope photographs of Figure 18 represent various modes of operation of the broadside antenna array. Chronologically, these are: At 2019, vertical-lobe, north-south polarized; at 2021, split-lobe, 30° to north and south of vertical, north-south polarized; at 2027, vertical-lobe, east-west polarized; and at 2030, split-lobe, 30° to east and west of vertical, east-west polarized. (Inasmuch as cross polarization was not employed in these observations, the fact that several polarizations were involved may be disregarded in interpreting the results.) The loss of the equatorial spread- F echo (that had been obtained in the height interval between 400 and 500 km in the first mode), upon using the second mode, indicates that the irregularities are aligned in the magnetic north-south direction. This is consistent with the retention of echoes on passing

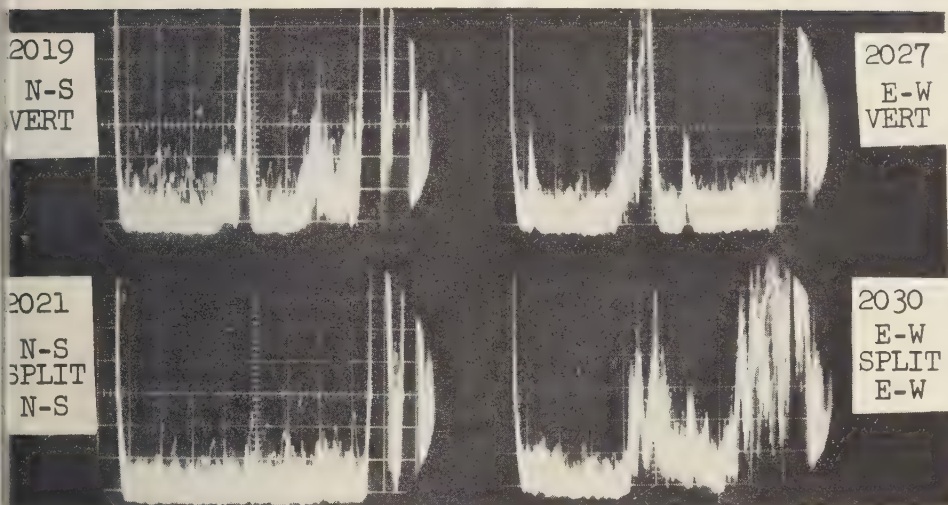


Fig. 18. Radar echoes at 8.4 Mc/s obtained at Huancayo, Peru, on November 17, 1959, using a broadside antenna beamed vertically or with lobes split 30° on either side of the vertical, with north and south or east and west splitting corresponding to the north-south or east-west polarization, respectively. (The reticule markers represent 100 km of height, measured positively to the right.)

from mode 3 to mode 4, since orthogonality on a field aligned model is still possible, for north-south oriented irregularities, throughout the east-west plane. Quantitatively, it can be concluded from the split-lobe observations only that the irregularities (measured along the magnetic lines of force) must be at least several wavelengths long at 8.4 Mc/s, i.e., longer than about 10 meters.

Results of the polarization-switching observations. Next is treated an *indirect* proof, based upon application of the magnetoionic theory, of the magnetic field alignment of the irregularities giving rise to equatorial spread F' . The reason that this additional proof is presented is that, as a consequence, a larger lower limit (hence a better estimate than was obtained above) is predicted for the length of the irregularities.

As far as magnetoionic considerations are concerned, the location near the magnetic equator of an antenna polarized as described above would, in principle, permit the launching of a single (quasi-transverse, 'QT,' or quasi-longitudinal, 'QL') magnetoionic component. However, the magnetic dip at Huancayo is about 2° , and the beamwidth of the antenna is appreciable. The fact that the broadside antenna array trans-

mits a significant amount of radiation at angles out to some 15° from the vertical may be interpreted on the magnetoionic theory as allowing the simultaneous propagation of both components. We can readily compute the angular departure to the north or south from orthogonality for which the ideal QT situation gives way to an admixture of the QL case, by using the inequalities [Ratcliffe, 1959]

$$\begin{array}{c} \text{QT} \\ \left(\frac{\sin^2 \theta}{2 \cos \theta} \right)^2 \left(\frac{\omega_H}{\omega} \right)^2 \gg \left(1 - \frac{\omega_N^2}{\omega^2} \right)^2 \\ \text{QL} \end{array}$$

for QT and QL, respectively. Here θ is the angle between the earth's magnetic field and the direction of propagation, ω is the exploring angular frequency, ω_H is the gyro-magnetic angular frequency, and ω_N is the plasma angular frequency. At the magnetic equator, for heights of the lower edge of the equatorial spread F' of about 300 km, and for the ionospheric situation of interest in Figure 19, we have $\omega/2\pi = 8.3$ Mc/s, $\omega_H/2\pi = 0.75$ Mc/s, and the plasma frequency in the neighborhood of the equatorial spread F' , $\omega_N/2\pi \lesssim 2$ Mc/s. Upon inserting these values in the inequality, it becomes

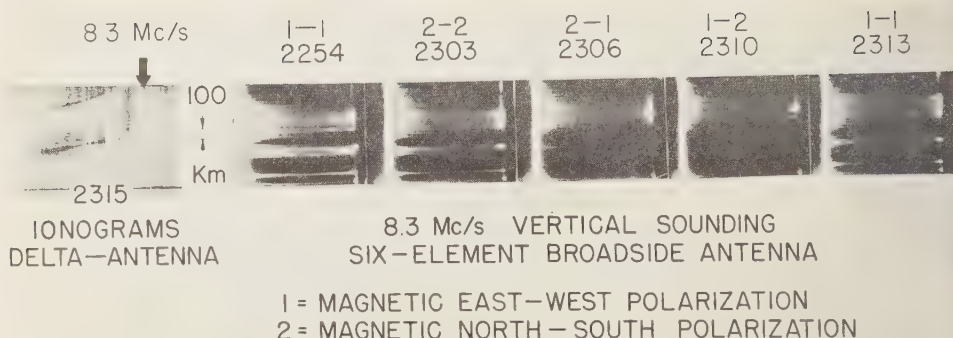


Fig. 19. Radar responses obtained at vertical incidence at a fixed frequency (8.3 Mc/s) using various combinations of polarization for transmitting and receiving with a broadside antenna array, compared with the associated sweep-frequency ionograms. Recorded at Huancayo, Peru, on December 20, 1958.

$$\begin{aligned} & \text{QT} \\ & \left(\frac{\sin^2 \theta}{2 \cos \theta} \right)^2 \gg \ll 120 \\ & \text{QL} \end{aligned}$$

and we derive the conditions

$$\frac{\sin^2 \theta}{2 \cos \theta} > 3 \sqrt{120}$$

or

$$\frac{\sin^2 \theta}{2 \cos \theta} < \frac{\sqrt{120}}{3}$$

for the QT and QL approximations, respectively. Thus, on inserting trial values for θ , we find that only out to 1° from exact transversality is the QT approximation strictly upheld. This result, in conjunction with the observations shown in Figure 16, will now be demonstrated to lead to some interesting consequences.

In Figure 19 are shown the experimental data obtained during the simultaneous occurrence at Huancayo of the equatorial and temperate-latitude varieties of spread F on December 20, 1958. The arrow adjacent to the ionogram corresponds to the exploring frequency of 8.3 Mc/s being used with the broadside array. A-scope (range versus amplitude) presentations of the echoes are displayed alongside the conventional ionogram. The height scales of the A-scope photographs are aligned with the height scale of the ionogram so that detailed comparisons can be made. We interpret the A-scope photographs

as indicating that the transition from the delta antenna to the broadside array has resulted in a considerable narrowing in the spread in time delays of the equatorial spread- F echoes, whereas there has been no appreciable change in the appearance of the temperate-latitude spread F . The interpretation of this narrowing, consistent with the slight pulse broadening in the oblique-scatter measurements, is that the scattering region must be comparatively thin.

On comparing the results of using the several polarization modes (Fig. 19), we note a fundamental difference between the behavior of the equatorial spread- F echoes and that of the temperate-latitude spread- F echoes; i.e., upon using cross polarization of the transmitting and receiving antennas, the equatorial spread- F echo disappears. This means that there has been no rotation of polarization during the propagation for otherwise an echo would have been observed. In view of the above computations, the fact that there is no rotation of polarization, and hence no echo, means that the QT situation must have been maintained. That is, orthogonality to the magnetic field lines must have been preserved for these echoes in order to avoid magnetoionic depolarizing effects. The only way that this QT condition can be maintained is that these echoes must arise from the equatorial spread F only in directions within 1° to north or south of the east-west plane of orthogonality, as we have seen from the above calculation. Thus, the angle of arrival of the equatorial spread- F echoes has been indirectly established, by the absence of

depolarization, as being a fan-shaped region with an aperture of about 2° , oriented magnetically east-west. (This region is tilted about 2° northward from the zenith, inasmuch as Huancayo is slightly north of the magnetic equator.) In order to produce this directional effect, it is required that the irregularities in the F region giving rise to equatorial spread F must be elongated along the earth's magnetic lines of force. Furthermore, it is inferred, from the 2° aperture consideration, that a lower limit for the extension of an irregularity along the magnetic field direction is 1000 meters.

Conclusions from the several observations of field alignment. It has been demonstrated in two ways that the irregularities are field aligned and that the spread in the time delays associated with equatorial spread F as seen on conventional ionograms must be attributed to echoes to the east and west of the zenith. We have discussed [Bowles and Cohen, 1957] the concept of magnetic field aligned irregularities giving rise to spread- F echoes. A model based upon such irregularities [Renau, 1960] considers the solid angle of arrival of spread F echoes for various magnetic latitudes, and in the equatorial case does, of course, lead to our experimental results.

From the time delays introduced on the ionograms, the patches of spread- F irregularities can be estimated to have a horizontal extent as great as 1000 km in the magnetic east-west direction. The actual thickness of the patches of scattering irregularities producing the equatorial spread F , as deduced from Figures 18 and 19, is of the order of 50 km, which is comparable to the scale height of the ionosphere at 300 km. This patch thickness is in agreement with the earlier estimate obtained from the pulse spread- F observed in oblique F scatter propagation.

DISCUSSION

Generality of these results. Despite the impression perhaps conveyed by the title of this article, the authors wish to point out that the present discussion of spread F in the equatorial ionosphere over Peru during the period of sunspot maximum is not necessarily generalizable to other longitudes along the magnetic equator, or to other phases of the sunspot cycle. Nevertheless, there is good reason to believe that the features of spread- F phenomena at the magnetic equator around the earth and throughout the

solar cycle will bear considerable resemblance to those described here, even though there may be differences in general details such as the nocturnal morphology.

Classification of spread- F configurations on equatorial ionograms. In view of the relative simplicity of the geometry of field aligned irregularities at the magnetic equator relative to the earth's surface, a study of the equatorial ionosphere can be expected to help resolve some of the phenomena due to these irregularities. There has heretofore been some question about the classification of configurations due to F -region scattering effects obtained on equatorial ionograms. A distinction has been suggested [McNicol, Webster, and Bowman, 1956] between 'range' and 'frequency' varieties of spread- F configurations appearing on ionograms in general, but usually the analysis of spread F has proceeded without the recognition of any real or apparent differences between the spread- F manifestations on conventional soundings. Also, ionograms are representations of spread- F phenomena that should not necessarily be interpreted at face value; e.g., the echoes may not originate from directly above the sounder.

For the classes of spread F appearing at the magnetic equator, the present experiments indicate an *observational* basis for a clear separation—whether the spread F does or does not support ionospheric scatter at 50 Mc/s. This is equivalent to determining whether irregularities are present comparable in size to the exploring wavelength, or smaller. Thus, whereas previously there has been a question whether a classification of spread- F forms on the ionograms was *justifiable*, we are now strongly compelled to distinguish between classes of spread F that do or do not propagate 50-Mc/s radio waves. The 'propagating' class of spread F must consist of irregularities aligned along the magnetic field lines and having at least one transverse dimension of the order of a wavelength at the exploring frequency, whereas the spread F that does not support scatter propagation must be viewed as produced by irregularities with *no* dimensions that small.

In respect to the nomenclature used in this paper for the spread- F echoes observed on equatorial ionograms, we have recognized a basic phenomenological distinction between two major classes of spread echoes by means of the ionospheric scattering effects just discussed. These

categories usually have the appearance of 'range spreading' and 'frequency spreading' on the ionograms. Since range spreading is characteristic of the equatorial region, we have used the term 'equatorial spread F ' to refer to it. Since the frequency spreading seen at the equator resembles that familiar in temperate latitudes, we have called it 'temperate-latitude spread F '. However, the use of this terminology does not imply that there is any *physical* resemblance of the temperate-latitude spread F observed at the equator to its counterpart in temperate latitudes, or of the equatorial spread F observed at the equator to similar phenomena observed at higher latitudes. Perhaps these points could be clarified by the use of a descriptive notation such as F_{eq} and F_{tq} for the basic categories of spread F in the equatorial ionosphere, by analogy with the notation E_{eq} used for q -type (equatorial) sporadic E . Then ' t ' in this system would designate 'temperate,' and the third subscript would represent the (temperate, polar, or equatorial) locale of observation.

We tentatively interpret the second class of spread F , the temperate-latitude variety, as arising from scattering, perhaps multiple, due to the presence of irregularities that are probably field aligned, that are large in all dimensions compared with the exploring wavelengths, but which effectively appear small in comparison with the large local wavelengths as the plasma frequency corresponding to the exploring wavelength is approached [Booker, 1955; Pitteway, 1960; Denisov, 1960]. Such an interpretation is consistent with the lack of oblique scattering at 50 Mc/s at times when temperate-latitude spread F is observable at Huancayo. Presumably, oblique scattering is absent because plasma resonances and the associated retardations cannot be excited at such a high frequency, and hence the large irregularities hypothesized would scatter only in the forward direction.

Comparison with F scatter in the Far East. Other VHF ionospheric forward-scatter studies [Bateman, Finney, Smith, Tveten, and Watts, 1959; Smith and Finney, 1960] have been conducted by the National Bureau of Standards in the Far East, and signals attributed to F scatter have been obtained. The paths over which these signals were propagated were about half as long as the 2580-km circuit discussed in this paper, implying a considerably larger scattering angle.

(Over our path of comparable length in Peru, the Arequipa \rightarrow Trujillo circuit, no spread- F signals were ever detected. However, the rhombic antennas used on that path strongly discriminated against scattering media above 120 km.) The signal strengths for the observed F scatter, in spite of the unfavorable comparison in scattering angles, were considerably greater in the Far East than near the equator. On the other hand, the appearance of the pulse delay recordings was very similar for the two sets of observations.

Transequatorial propagation. Besides the earlier amateur activities in this field [Ferrell, 1951], there was a more elaborate program [Southworth, 1960] by radio amateurs during the IGY for studying long-range, transequatorial propagation at VHF. The mechanism of this propagation now appears to be explicable from the results of recent [Dueño, 1960; Villard, Stein, and Yeh, 1957] backscatter studies. The generally accepted interpretation is that during years of high sunspot activity the nighttime equatorial ionosphere has a dome-shaped configuration that is conducive, by a refraction mechanism, to propagating rays in so-called 'supermode' trajectories that do not return to earth at the path midpoint. It is our impression that the rapidly fluctuating signals that occur from time to time over these transequatorial paths during spread- F conditions near the equator are attributable to an admixture of forward-scattering effects (superimposed on traversing the equatorial region), but that the controlling means of propagation is always refraction.

The origin of spread F . The time variation of spread- F occurrences has been shown in Figures 7, 8, and 9, and it is not readily explicable in terms of some geophysical agency that creates spread echoes on some occasions and not on others. In the equatorial ionosphere there has been established a generally negative correlation of spread- F activity with magnetic-disturbance indices [Lyon, Skinner, and Wright, 1958]. Although this same correlation may be noted in the data of Figures 7, 8, and 9, it does not seem to fulfill the desideratum of suggesting the causative factor.

A correlation has been observed [Booker and Wells, 1938; Osborne, 1951] between the rise after sunset of the equatorial F layer and the occurrence of spread F . A current theoretical treatment [Martyn, 1959] of this phenomenon

proposes a causal relationship between the rising of the F layer and the formation of spread- F irregularities. This theory would also predict that a falling layer would *not* generate the irregularities. However, the observations of spread F indicate that it can occur during the rising and/or falling phases of F -layer displacements.

The following discussion is a plausibility argument toward understanding the occurrence or nonoccurrence of spread F . This discussion leads to an hypothesis whose predictions can be compared with experimental observations.

The criterion for obtaining radar echoes from magnetic field aligned irregularities is that the exploring radiation must be directed orthogonally to the lines of force. Anisotropic electron irregularities, the most general case, are defined as distributions of ionization having a large autocorrelation distance as their longest dimension and small autocorrelation distances as their shorter dimensions. Stated differently, these irregularities have their highest ionization gradients transverse to their longest dimensions, so that magnetic field aligned irregularities would necessarily be characterized by steep gradients transverse to the magnetic field lines.

But this is precisely the property of a distribution of ionization having its isoionic contours aligned with the magnetic field; that is, the steepest realizable ionization gradients, which are those normal to the contours, would in this case be oriented relative to the magnetic field exactly as those for field aligned irregularities. Such a configuration of mean isoionic contours parallel to the magnetic field lines would have maximum stability relative to other configurations, since there would be no tendency for diffusion *along* the magnetic lines of force, and diffusion *across* the lines of force would be inhibited. If mean isoionic contours oriented in this way were to give rise to field aligned irregularities, the irregularities themselves, in turn, would partake of this stability. We should like to hypothesize that the presence of electron density contours parallel to the earth's magnetic field lines is a necessary condition for the *formation* as well as the stability of field aligned irregularities in the equatorial F region. The actual mechanism *sufficient* for the formation of irregularities is not yet understood, but we postulate that the proposed contour orientation

is *necessary* for that process. Then, by studying mean electron density contours and their orientation relative to the magnetic field, we can predict on this model when the production of irregularities appropriately situated (i.e., located along the locus of orthogonality to the magnetic field lines) to support radio wave scattering from a given station is *possible*. These results can then be compared with the observed scattering effects in the F region.

Only a limited number of electron density profiles along the Huancayo meridian plane have been examined up to now for evidence with which to test the present hypothesis. The data so far available are generally consistent with the hypothesis. An example of this agreement is presented in Figure 20. Each of the three components of this figure is a superposition of the magnetic lines of force, the electron density profiles, and the loci of orthogonality to the lines of force from the several stations indicated. On the evening of November 18, 1958, there was no spread F at Huancayo at 2000, but there was equatorial spread F at Huancayo on November 19, 1958, at that time, but not at 2200. This is in accordance with what we would predict from Figure 20 on the basis of the above discussion. On November 18, only for Chimbote was there some possibility of elevated spread F at 2000, but none was actually observed at any station. At 2000 on November 19 there is a suggestion that there may possibly be spread F at Chimbote, which does occur, but rather weakly. At 2200 there is some expectation of observing spread F at Chimbote, Chiclayo, and perhaps Huancayo, but it is not realized. Thus, in this example, spread F frequently occurs when it is predicted on our hypothesis, and it does not occur unexpectedly.

It should be explained that the lines of force in Figure 20 are those for the earth's dipole field, but centered about the magnetic equator. Also, the true-height data for the various locations have been corrected in time according to the longitudinal separations from Huancayo. The nearly straight lines emanating from each station are the loci of orthogonality. (Right angles are not, in general, preserved in Fig. 20, owing to the vertical exaggeration in scale.) A reasonable criterion for predicting the possibility of spread F on the basis of the present hypothesis is that the parallelism of the lines of force with the

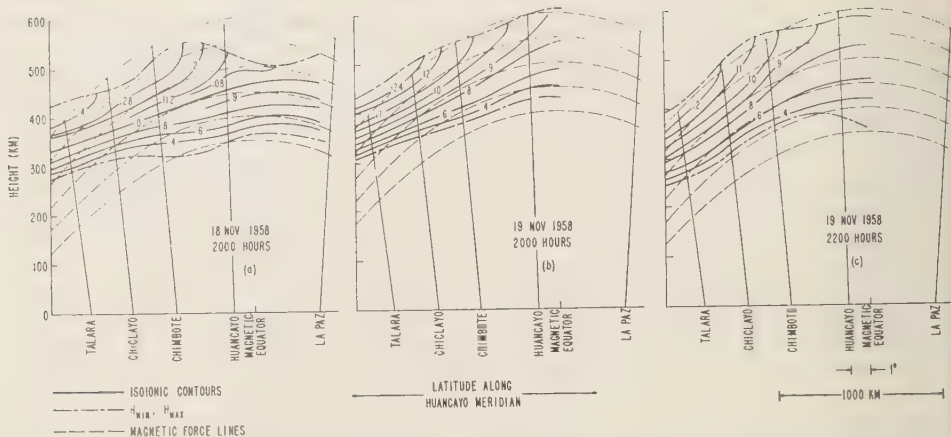


Fig. 20 Comparison of mean electron density contours along the Huancayo meridian with the earth's magnetic lines of force. (The numbers associated with the contours refer to the plasma frequency in megacycles per second.)

mean isoionic contours should extend over some kilometers, and not simply in the vicinity of a point of tangency. Inasmuch as equatorial spread F usually occurs somewhat below the minimum height of the F layer, the prediction procedure suggested for this variety of spread echo is to utilize the contour for h_{\min} of the F layer, or an extrapolated contour, parallel to and just beneath that for h_{\min} .

In view of the absence of spread F during the day and its presence at night, the fact that the electron density contours near the equator are fairly horizontal by day but arched at night with curvatures similar to those of the lines of force lends credence to the suggested explanation. It is difficult to understand how the proposed condition for spread- F formation could apply at latitudes where the magnetic field departs appreciably from horizontality, since at such latitudes there would be less likelihood of realizing the geometrical requirements. If magnetic disturbances at higher latitudes produce sufficiently great changes in slope of the mean isoionic contours, however, the present hypothesis may be extensible to those latitudes. Then the reversal in sign (above 20° magnetic latitude) of the correlation of spread- F echoes with magnetic activity [Lyon, Skinner, and Wright, 1958] could be understood, since the distortions of the normal contour configurations produced by magnetic activity would enhance the possibility of spread

F above that latitude and diminish it near the equator.

CONCLUSIONS

Experimental evidence has been presented (1) that a class of spread F identifiable on equatorial ionograms as equatorial spread F is capable of supporting F scatter propagation at 50 Mc/s; (2) that another class identifiable as temperate-latitude spread F does not have this property; (3) that the equatorial spread- F time delays are not introduced (as had been suggested [Booker, 1956]) by a forward-scattering screen situated below the F layer; (4) that the irregularities giving rise to equatorial spread F are magnetic field aligned; (5) that the irregularities are 1 km or longer in the direction of the magnetic field; (6) that the irregularities measured in one (or possibly both) dimension(s) transverse to the magnetic field are of width 10 meters or less (the existence of structure as small as this at, e.g., 300 km, is to be contrasted with the ionospheric parameters at that height, viz., a mean free path for electrons of about 150 meters, a scale height of 50 km, an electron gyroradius of the order of 10 cm, an ion gyroradius (for O^+) of the order of 20 meters); (7) that the spread in the ionosonde equatorial spread- F time delays arises from echoes arriving over a wide range of elevation angles in the magnetic east-west plane; (8) that these echoes result from scattering by

comparatively thin sheets or patches of irregularities whose thickness is of the order of 50 km; (9) that these patches sometimes extend horizontally in the magnetic east-west direction over distances of the order of 1000 km; (10) that the scattering layer in the F region is at about the lowest height of the equatorial spread F observed on the ionograms; (11) that this layer can occur at the bottom of the F layer or as much as 100 km below it; (12) that the scattering layer can occur up to heights of 450 km or more.

In the discussion, the following points have been made: (1) It is submitted that two kinds of spread F are identifiable on equatorial ionograms, designated in this paper as equatorial and temperate-latitude spread F , and that they can be distinguished observationally by their ability or failure, respectively, to support oblique-incidence ionospheric scattering at 50 Mc/s. (2) It is suggested that a necessary condition for the occurrence of spread F in the equatorial ionosphere is that the surfaces of constant electron density be approximately parallel to the magnetic field lines.

Acknowledgments. Individuals contributing to this research included, for the National Bureau of Standards, R. C. Kirby, B. B. Balsley, K. W. Sullivan, P. G. Sulzer, H. G. Sellery, P. P. Viezbicke, W. H. Harding, W. H. Daniels, W. Calvert, J. T. Brown, Jr., C. E. Hoff, W. M. Rogers, F. B. Haller, C. Y. Matsunaka, G. L. McConnell, J. L. Valega, N. W. Bivens, R. W. Lester, G. W. Pickering, Jean Lawrence Bryant, Alice Jenkins Ochs, Augusta H. O'Brien, Virginia C. Bartow, J. R. DeLine, A. O. Crawley, M. C. Weeg, B. A. Benway, O. R. Dallman, and many others. Appreciation is expressed to G. H. Stonehocker and J. W. Wright of NBS for the true-height calculations utilized in this paper. We are grateful for the cooperation of the U. S. Department of State and of the IGY Committees and Governments of Argentina, Brazil, Chile, Ecuador, and Peru. The generous collaboration of the Instituto Geofísico de Huancayo was of invaluable assistance. We are particularly indebted to the following organizations and fellow Americans. *Of Argentina:* Armada República Argentina, Fuerza Aérea Argentina, Lt. José Rodríguez (ARA), Héctor Cáffaro, Comandante Eloy del Valle Castro of the Gendarmería Nacional República Argentina and his assistants, Auxiliar Norberto Carlos García, and Ayudante Vicente Spatazza. *Of Brazil:* Professor L. deQ. Orsini, Universidade de São Paulo. *Of Chile:* Coronel Alberto Stegmaier, Padre Germán Saa, The Antofagasta (Chili) and Bolivia Railway Company, Ltd. *Of Ecuador:* Armada del Ecuador, Lt. Fernando García of the Ejército Ecuatoriano. *Of Mexico:* Geoff. W. Lord, XEIGE, XEIGL. *Of Peru:* Alberto A. Giesecke, Jr.,

Director-Técnico, Instituto Geofísico de Huancayo, and his colleagues, Max Tábara, Rómulo Alvarez, Jorge Ghersi. Finally, we are appreciative of the generosity of the several Latin American governments in the assignment of radio frequencies and the removal of all customs barriers to our experimental activities.

Support for this research was provided by the U. S. National Committee for the IGY through a grant from the National Science Foundation, by the Voice of America (U. S. Information Agency), and by funds of the National Bureau of Standards.

REFERENCES

- Bailey, D. K., R. Bateman, and R. C. Kirby, Radio transmission at VHF by scattering and other processes in the lower ionosphere, *Proc. IRE*, **43**, 1181-1231, 1955.
- Bateman, R., J. W. Finney, E. K. Smith, L. H. Tveten, and J. M. Watts, IGY observations of F -layer scatter in the Far East, *J. Geophys. Research*, **64**, 403-405, 1959.
- Booker, H. G., On the level at which fading is imposed on waves reflected vertically from the ionosphere, *J. Atmospheric and Terrest. Phys.*, **7**, 343-344, 1955.
- Booker, H. G., Turbulence in the ionosphere with applications to meteor-trails, radio-star scintillation, auroral radar echoes, and other phenomena, *J. Geophys. Research*, **61**, 673-705 (esp. p. 704), 1956.
- Booker, H. G., and H. W. Wells, Scattering of radio waves by the F region of the ionosphere, *Terrest. Magnetism and Atmospheric Elec.*, **43**, 249-256, 1938.
- Bowles, Kenneth, and Robert Cohen, NBS equatorial region VHF scatter research program for the IGY, *QST*, **41**, 11-15, 1957.
- Bowman, G. G., Further studies of 'spread- F ' at Brisbane, I, *Planetary and Space Sci.*, **2**, 133-149, 1960.
- Briggs, B. H., A study of the ionospheric irregularities which cause spread- F echoes and scintillations of radio stars, *J. Atmospheric and Terrest. Phys.*, **12**, 34-45, 1958.
- Calvert, Wynne, and Robert Cohen, The interpretation and synthesis of certain spread- F configurations appearing on equatorial ionograms, *J. Geophys. Research*, **66** (in press), 1961.
- Dagg, M., Diurnal variations of radio-star scintillations, spread F , and geomagnetic activity, *J. Atmospheric and Terrest. Phys.*, **10**, 204-214, 1957.
- Denisov, N. G., On the influence of the reflection region on radio-wave scattering in the ionosphere, *Izvest. VUZ, Radiofizika*, **3**, 208-215, 1960.
- Dueño, Braulio, Peculiarities and seasonal variations of transequatorial backscatter echoes as observed at Mayagüez, Puerto Rico, *J. Geophys. Research*, **65**, 1691-1698, 1960.
- Ferrell, O. P., (Abstract) Very-high-frequency propagation in the equatorial region, *Proc. IRE*, **39**, 719, 1951.

- Glover, F. N., A survey of spread F , *NBS Tech. Note 82*, November 1960.
- Knecht, R. W., Possibility of detecting ionospheric drifts from the occurrence of spread F echoes at low latitudes, *Nature*, 187, 927, 1960.
- Lyon, A. J., N. J. Skinner, and R. W. Wright, Equatorial spread F and magnetic activity, *Nature*, 181, 1724-1725, 1958.
- McNicol, R. W. E., H. C. Webster, and G. G. Bowman, A study of Spread- F ionospheric echoes at night at Brisbane, 1, Range spreading (experimental), *Australian J. Phys.*, 9, 247-271, 1956.
- Martyn, D. F., The normal F region of the ionosphere, *Proc. IRE*, 47, 147-155, 1959.
- Osborne, B. W., Ionospheric behaviour in the F_2 region at Singapore, *J. Atmospheric and Terrest. Phys.*, 2, 66-78, 1951.
- Pitteway, M. L. V., The reflexion of radio waves from a stratified ionosphere modified by weak irregularities, *Proc. Roy. Soc. London, A*, 254, 86-100, 1960.
- Ratcliffe, J. A., *The Magneto-Ionic Theory and Its Applications to the Ionosphere*, Cambridge University Press, 206 pp., 1959.
- Reber, G., World-wide spread F , *J. Geophys. Research*, 61, 157-164, 1956.
- Renau, Jacques, Theory of spread F based on aspect-sensitive back-scattered echoes, *J. Geophys. Research*, 65, 2269-2277, 1960.
- Smith, E. K., Jr., and James W. Finney, Peculiarities of the ionosphere in the Far East: A report on IGY observations of sporadic E and F -region scatter, *J. Geophys. Research*, 65, 885-892, 1960.
- Southworth, M. P., Nighttime equatorial propagation at 50 Mc/s, First results from an IGY amateur observing program, *J. Geophys. Research*, 65, 601-607, 1960.
- Villard, O. G., Jr., Sidney Stein, and K. C. Yeh, Studies of transequatorial ionospheric propagation by the scatter-sounding method, *J. Geophys. Research*, 62, 399-412, 1957.
- Wells, H. W., F scatter at Huancayo, Peru, and its relation to radio-star scintillations, *J. Geophys. Research*, 59, 273-277, 1954.
- Wright, R. W., J. R. Koster, and N. J. Skinner, Spread F -layer echoes and radio-star scintillation, *J. Atmospheric and Terrest. Phys.*, 8, 240-246, 1956.
- Wright, R. W., and N. J. Skinner, Equatorial spread F , *J. Atmospheric and Terrest. Phys.*, 15, 121-125, 1959.

(Manuscript received January 14, 1961.)

Doppler Shifts and Faraday Rotation of Radio Signals in a Time-Varying, Inhomogeneous Ionosphere

Part II. Two-Signal Case

JOHN M. KELSO

*Space Technology Laboratories, Inc.
Los Angeles 45, California*

Abstract. Following procedures applied in Part I [Kelso, 1960a] to a single radio signal transmitted to the ground from a space vehicle in or above the ionosphere, equations are obtained in Part II for the time rate of change of the phase difference between two signals whose refractive indices are slightly different. These results are applied to the frequency shifts experienced in a two-frequency Doppler experiment of the Seddon-Jackson type, and to the determination of the rate of Faraday rotation of the plane of polarization.

The principal restriction on the generality of the results is that the ionosphere is treated as quasi-isotropic; i.e., the ray paths are obtained by methods that would be exact in a slowly varying isotropic medium, but the refractive index is permitted to be a function of ray direction (implying an anisotropic ionosphere).

The following conditions prevail: (a) the (slowly varying) ionosphere may be a general function of three spatial co-ordinates and of time; (b) the vehicle may follow an arbitrary (nonrelativistic) trajectory; (c) the magnetic field, which characterizes the dependence of the refractive index on direction, may have arbitrary form.

1. *Introduction.* A previous paper (denoted Part I) by the author [Kelso, 1960a] derived expressions for the frequency (or phase) shift suffered by a single radio signal transmitted to the ground from a space vehicle traveling in or above a time-varying, nonuniform, and anisotropic ionosphere. In the present paper, we wish to derive corresponding results for the time-variation of the phase difference between two simultaneous signals whose refractive indices are slightly different. This result will then be applied to the Seddon-Jackson two-frequency Doppler experiment for the measurement of electron density [Seddon, 1953; Jackson and Seddon, 1958] and to the Faraday rotation of the plane of polarization [Browne, Evans, Hargreaves, and Murray, 1956; Daniels and Bauer, 1959].

As in Part I, the problem studied has the following general properties: (1) the (slowly-varying) refractive index is an arbitrary function of position and time; (2) the vehicle moves in an arbitrary (nonrelativistic) trajectory; (3) the magnetic field may have an arbitrary form; (4) the effects caused by variations in the shape of the ray are included; (5) the ray paths are based on quasi-isotropic calculations, i.e., the methods used are exact for an isotropic medium,

but the refractive index is treated as a function of direction.

The basic notation used below is essentially the same as that introduced in Part I:

- x, y, z = positional co-ordinates.
- ξ, η, ζ = vehicle co-ordinates.
- t = time.
- x', y' = $\partial x/\partial z, \partial y/\partial z$ = ray slopes.
- Γ = denotes a ray path.
- ds = element of arc along a ray.
- P = phase path (meter).
- ϕ = phase (radians) = $k_0 P$.
- k_0 = $(2\pi f)/c$ (meter⁻¹).
- f = frequency (sec⁻¹).
- c = velocity of light in free space.
- V_{11} = component of vehicle velocity along a ray path.
- N = electron density (meter⁻³).

2. *Variation of phase path differences between two signals.* In the present section, we wish to compute the variation in the difference in phase path between two signals whose refractive indices are slightly different. The medium is assumed to be quasi-isotropic. The procedure is a generalization of a method [Altshuler, unpublished memorandum] used to derive the equivalent of

constant. Using subscripts to denote partial differentiation, we have:

$$\begin{aligned}
 \delta(P_2 - P_1) = & \int_0^{t_2} [g + g_x \nabla x + g_{x'} \nabla x' \\
 & + g_t \delta t + g_\epsilon \delta \epsilon + \frac{1}{2} g_{xx} (\nabla x)^2 \\
 & + \frac{1}{2} g_{x'x'} (\nabla x')^2 \\
 & + \frac{1}{2} g_{tt} (\delta t)^2 + \frac{1}{2} g_{\epsilon\epsilon} (\delta \epsilon)^2 + g_{xx'} \nabla x \nabla x' \\
 & + g_{xt} \nabla x \delta t + g_{x\epsilon} \nabla x \delta \epsilon \\
 & + g_{x't} \nabla x' \delta t + g_{x'\epsilon} \nabla x' \delta \epsilon \\
 & + g_{t\epsilon} \delta t \delta \epsilon] dz \\
 & - \int_0^{t_1} [g + g_x \Delta x + g_{x'} \Delta x' + g_t \delta t \\
 & + \frac{1}{2} g_{xx} (\Delta x)^2 + \frac{1}{2} g_{x'x'} (\Delta x')^2 \\
 & + \frac{1}{2} g_{tt} (\delta t)^2 + g_{xx'} \Delta x \Delta x' \\
 & + g_{xt} \Delta x \delta t + g_{x't} \Delta x' \delta t] dz \\
 & - \int_0^{t_1} [g + g_x \delta x + g_{x'} \delta x' + g_\epsilon \delta \epsilon \\
 & + \frac{1}{2} g_{xx} (\delta x)^2 + \frac{1}{2} g_{x'x'} (\delta x')^2 \\
 & + \frac{1}{2} g_{\epsilon\epsilon} (\delta \epsilon)^2 + g_{xx'} \delta x \delta x' \\
 & + g_{x\epsilon} \delta x \delta \epsilon + g_{x'\epsilon} \delta x' \delta \epsilon] dz \\
 & + \int_0^{t_1} g dz \quad (3)
 \end{aligned}$$

Let us define new quantities Dx and Dx' from the following equations:

$$\begin{aligned}
 \nabla x &= Dx + \delta x \\
 \nabla x' &= Dx' + \delta x' \quad (4)
 \end{aligned}$$

Replace ∇x and $\nabla x'$ in equation 3 by the expressions in equation 4. Now, we note that Δx and $\Delta x'$ are differences between the x position and the slope of the two rays of type 1 for times t and $t + \delta t$. Hence, Δx and $\Delta x'$ are of order δt . Similarly, Dx and Dx' are corresponding differences for the two rays of type 2 at times t and $t + \delta t$. Hence, Dx and Dx' are also of order δt . On the other hand, δx , $\delta x'$, and $\delta \epsilon$ do not vanish as $\delta t \rightarrow 0$. In what follows, we will neglect terms of order $(\delta t)^2$. Thus, we now discard squares and cross products of Δx , $\Delta x'$,

Dx , Dx' , and δt . Then, after some rearrangement, equation 3 may be written

$$\begin{aligned}
 \delta(P_2 - P_1) = & \int_0^{t_2} (g_x Dx + g_{x'} Dx') dz \\
 & - \int_0^{t_2} (g_x \Delta x + g_{x'} \Delta x) dz \\
 & + \int_{t_1}^{t_2} g_\epsilon \delta \epsilon dz + \int_0^{t_2} (g_{x\epsilon} Dx \\
 & + g_{x'\epsilon} Dx' + g_{t\epsilon} \delta t) \delta \epsilon dz \\
 & + \int_0^{t_2} [g_{xx} Dx \delta x + g_{x'x'} Dx' \delta x' \\
 & + g_{xx'} (Dx \delta x' + Dx' \delta x) \\
 & + g_{xt} \delta x \delta t + g_{x't} \delta x' \delta t] dz \\
 & + \int_{t_1}^{t_2} [g_x \delta x + g_{x'} \delta x' + \frac{1}{2} g_{xx} (\delta x)^2 \\
 & + \frac{1}{2} g_{x'x'} (\delta x')^2 + \frac{1}{2} g_{\epsilon\epsilon} (\delta \epsilon)^2 + g_{xx'} \delta x \delta x' \\
 & + g_{x\epsilon} \delta x \delta \epsilon + g_{x'\epsilon} \delta x' \delta \epsilon] dz \quad (5)
 \end{aligned}$$

Since the first two integrals each represent variations in a ray, along which Euler's equation applies, we may use an analysis similar to that of Part I to show that the two integrals reduce to

$$\begin{aligned}
 & \int_0^{t_2} (g_x Dx + g_{x'} Dx') dz \\
 & - \int_0^{t_2} (g_x \Delta x + g_{x'} \Delta x) dz \\
 & = [(Dx - \Delta x) g_{x'}]_{z=t_1}
 \end{aligned}$$

Now, the difference $Dx - \Delta x$ at $z = t_2$ is of order $(\delta t)^2$, and hence we neglect the first two integrals in equation 5.

Equation 5 also contains two integrals which we may write in the form

$$\int_{t_1}^{t_2} Q dz = (Q \delta t)_{z=t_2} + \text{terms of order } (\delta t)^2$$

Then, dividing both sides of equation 5 by δt , letting $\delta t \rightarrow 0$, and writing

$$\lim_{\delta t \rightarrow 0} \frac{Dx}{\delta t} = \frac{\partial x}{\partial t}$$

$$\lim_{\delta t \rightarrow 0} \frac{Dx'}{\delta t} = \frac{\partial x'}{\partial t}$$

we get

$$\begin{aligned}
 \frac{d(P_2 - P_1)}{dt} &= \left(g_{\epsilon} \delta\epsilon \frac{d\zeta}{dt} \right)_{z=\zeta_2} \\
 &+ \int_0^{\zeta_2} \left(g_{x\epsilon} \frac{\partial x}{\partial t} + g_{x'\epsilon} \frac{\partial x'}{\partial t} + g_{\epsilon\epsilon} \right) \delta\epsilon \, dz \\
 &+ \int_0^{\zeta_2} \left[g_{xx} \delta x \frac{\partial x}{\partial t} + g_{x'x'} \delta x' \frac{\partial x'}{\partial t} \right. \\
 &+ g_{xx'} \left(\frac{\partial x}{\partial t} \delta x' + \frac{\partial x'}{\partial t} \delta x \right) \\
 &+ g_{x\epsilon} \delta x + g_{x'\epsilon} \delta x' \left. \right] dz \\
 &+ \left\{ [g_{xx} \delta x + g_{x'x'} \delta x' + \frac{1}{2} g_{xx} (\delta x)^2 \right. \\
 &+ \frac{1}{2} g_{x'x'} (\delta x')^2 + \frac{1}{2} g_{\epsilon\epsilon} (\delta\epsilon)^2 + g_{xx'} \delta x \delta x' \\
 &+ g_{x\epsilon} \delta x \delta\epsilon + g_{x'\epsilon} \delta x' \delta\epsilon \left. \right] \frac{d\zeta}{dt} \Big\}_{z=\zeta_2} \quad (6)
 \end{aligned}$$

Since δx vanishes at $z = \zeta_2$ in the limit as $\delta t \rightarrow 0$, the last term reduces to

$$\left\{ [g_{x'x'} \delta x' + \frac{1}{2} g_{x'x'} \delta x' \delta\epsilon + g_{\epsilon\epsilon} (\delta\epsilon)^2] \frac{d\zeta}{dt} \right\}_{z=\zeta_2},$$

where we neglect a second-order term in $(\delta x')^2$. In the two special cases considered below, we find that $g_{\epsilon\epsilon} \equiv 0$. Hence, equation 6 reduces to

$$\begin{aligned}
 \frac{d(P_2 - P_1)}{dt} &= \left(g_{\epsilon} \delta\epsilon \frac{d\zeta}{dt} \right)_{z=\zeta_2} \\
 &+ \int_0^{\zeta_2} \left(g_{x\epsilon} \frac{\partial x}{\partial t} + g_{x'\epsilon} \frac{\partial x'}{\partial t} + g_{\epsilon\epsilon} \right) \delta\epsilon \, dz \\
 &+ \int_0^{\zeta_2} \left[g_{xx} \delta x \frac{\partial x}{\partial t} + g_{x'x'} \delta x' \frac{\partial x'}{\partial t} \right. \\
 &+ g_{xx'} \left(\frac{\partial x}{\partial t} \delta x' + \frac{\partial x'}{\partial t} \delta x \right) \\
 &+ g_{x\epsilon} \delta x + g_{x'\epsilon} \delta x' \left. \right] dz \\
 &+ \left\{ [g_{x'x'} \delta x' + g_{x'\epsilon} \delta x' \delta\epsilon] \frac{d\zeta}{dt} \right\}_{z=\zeta_2} \quad (7)
 \end{aligned}$$

The second integral depends on δx and $\delta x'$ which are differences in position and slope of the two rays corresponding to the two different refractive indices. Generally, in the problems of interest, these differences are neglected. This

integral provides a means for computing the error incurred in any specific problem by assuming both signals to travel a common path.

The preceding analysis makes it clear that the results may be extended to the three-dimensional case without changes in the procedure. Thus, for the case where the two signals are assumed to travel a common path, we may write:

$$\begin{aligned}
 \frac{d(P_2 - P_1)}{dt} &= \left(g_{\epsilon} \delta\epsilon \frac{d\zeta}{dt} \right)_{z=\zeta} \\
 &+ \left\{ [g_{x'} \delta x' + g_{x'\epsilon} \delta x' \delta\epsilon] \frac{d\zeta}{dt} \right\}_{z=\zeta} \\
 &+ \left\{ [g_{y'} \delta y' + g_{y'\epsilon} \delta y' \delta\epsilon] \frac{d\zeta}{dt} \right\}_{z=\zeta} \\
 &+ \int_0^{\zeta} \left(g_{x\epsilon} \frac{\partial x}{\partial t} + g_{x'\epsilon} \frac{\partial x'}{\partial t} \right. \\
 &+ g_{y\epsilon} \frac{\partial y}{\partial t} + g_{y'\epsilon} \frac{\partial y'}{\partial t} + g_{\epsilon\epsilon} \left. \right) \delta\epsilon \, dz \quad (8)
 \end{aligned}$$

where we have dropped the subscript on ζ , but where we have retained $\delta x'$ evaluated at the vehicle in the second and third terms.

At the beginning of this analysis we arbitrarily terminated the Taylor expansion at terms of the second order. If higher order terms were included, the result given by equation 8 would not be changed, because higher order terms in δt , Dx , Dx' , Δx , and $\Delta x'$ all vanish as $\delta t \rightarrow 0$, and higher order terms in $\delta\epsilon$ vanish for the special cases treated here. However, the general result corresponding to equation 7 would be changed, because we would obtain additional integrals containing higher order terms in δx and $\delta x'$. The use of such high-order terms might be required for a problem where the two signals follow quite different paths; but this type of problem will not be treated here.

3. Doppler experiment. In a two-frequency Doppler experiment of the type carried out by Seddon and Jackson, the vehicle transmits continuously and coherently on a frequency f_0 and on a second frequency qf_0 , where q is a precisely maintained constant. The electron density may be computed from a measurement on the ground of the function $d/dt (P_2 - P_1)$. We wish first to relate this function to the instantaneous frequencies received on the ground.

The phase ϕ of a signal is equal to k_0 times the phase path P , where $k_0 = 2\pi f_0/c$, c being the

velocity of light in free space and f_0 the transmitted frequency. We write,

$$\phi_1 = (2\pi f_0/c)P_1$$

$$\phi_2 = (2\pi qf_0/c)P_2$$

Then,

$$\begin{aligned} \frac{d}{dt}(P_2 - P_1) &= \frac{c}{2\pi f_0} \frac{d}{dt} \left(\frac{\phi_2}{q} - \phi_1 \right) \\ &= \frac{c}{2\pi qf_0} \frac{d}{dt} (\phi_2 - q\phi_1) \end{aligned} \quad (9)$$

The instantaneous frequency f is given by

$$f = \frac{1}{2\pi} \frac{d\phi}{dt}$$

Hence, equation 9 may also be written

$$\frac{d}{dt}(P_2 - P_1) = \frac{c}{f_0 q} (f_2 - qf_1)$$

where f_1 is the instantaneous received frequency of the signal transmitted at the frequency f_0 , and f_2 is the instantaneous received frequency of the signal transmitted at qf_0 . We assume that the frequencies are substantially constant over a long enough time to permit a measurement of the frequency difference $f_2 - qf_1$. We wish now to use the results of the preceding section to show that this measurement may be used, under suitable conditions, to calculate the electron density in the vicinity of the vehicle.

The square of the refractive index μ is written as

$$\mu^2 = 1 - \frac{2KN}{f^2} \quad (10)$$

where, using rationalized MKS units,

N = electron density in electrons per cubic meter.

$K = e^2/(8\pi^2 m \epsilon_0) \doteq 40 \text{ meter}^3 \text{ sec}^{-2}$.

e = charge of an electron = 1.602×10^{-19} coulomb.

m = mass of an electron = 9.108×10^{-31} kilogram.

ϵ_0 = Permittivity of free space = 8.85×10^{-12} farad meter $^{-1}$.

f = frequency in cycles per second.

Equation 10 corresponds to the case of the ordinary mode in quasi-transverse propagation and its use is equivalent to treating the medium as isotropic. For sufficiently high frequency, this

expression is applicable over a wide range of ionospheric conditions. (See, for example, *Mitra*, [1952] or *Ratcliffe*, [1959]).

Since $2KN/f^2$ is small compared with unity, μ may be obtained approximately from equation 10 as

$$\mu \doteq 1 - \frac{KN}{f^2}$$

We now write $1/f^2 = \epsilon/f_0^2$. Then we obtain μ_1 for signal 1 when $\epsilon_1 = 1$, and μ_2 for signal 2 when $\epsilon_2 = \epsilon_1 + \delta\epsilon = 1/q^2$. Then $\delta\epsilon = (1/q^2 - 1)$. Thus,

$$\mu = 1 - \frac{\epsilon KN}{f_0^2}$$

We note that N is a function of position and time.

Although the equations obtained in the preceding section may be used for a quasi-isotropic medium in which μ is a function of ray direction, we restrict our present attention to the isotropic case because of the necessity to introduce ϵ in a convenient manner.

The results of section 2 require consideration of the function

$$\begin{aligned} g(x, x', y, y', z, t, \epsilon) \\ = \mu(x, y, z, t, \epsilon) \sqrt{1 + \left(\frac{\partial x}{\partial z}\right)^2 + \left(\frac{\partial y}{\partial z}\right)^2} \end{aligned}$$

Since the radical is not a function of ϵ , we may write

$$\begin{aligned} \frac{\partial g}{\partial \epsilon} \delta\epsilon &= \frac{\partial \mu}{\partial \epsilon} \delta\epsilon \sqrt{1 + x'^2 + y'^2} \\ &= \frac{-K}{f_0^2} (1/q^2 - 1) N \sqrt{1 + x'^2 + y'^2} \end{aligned} \quad (11)$$

Equation 8 may be written,

$$\begin{aligned} \frac{d(P_2 - P_1)}{dt} &= \left(\frac{\partial \mu}{\partial \epsilon} \delta\epsilon \sqrt{1 + x'^2 + y'^2} \frac{d\xi}{dt} \right)_{z=z_1} \\ &+ \left[\mu_2 \frac{d\xi}{dt} \delta(\sec x) \right]_{z=z_1} \\ &+ \int_0^z \frac{\partial}{\partial \epsilon} \left(\frac{\partial g}{\partial x} \frac{\partial x}{\partial t} + \frac{\partial g}{\partial x'} \frac{\partial x'}{\partial t} \right. \\ &+ \left. \frac{\partial g}{\partial y} \frac{\partial y}{\partial t} + \frac{\partial g}{\partial y'} \frac{\partial y'}{\partial t} \right) \delta\epsilon dz \\ &+ \int_0^z \frac{\partial}{\partial t} \left(\frac{\partial g}{\partial \epsilon} \right) \delta\epsilon dz \end{aligned} \quad (12)$$

where

$$\delta \sec x = \left[\delta x' \frac{\partial}{\partial x'} \sqrt{1 + (x')^2 + (y')^2} + \delta y' \frac{\partial}{\partial y'} \sqrt{1 + (y')^2 + (y')^2} \right] \quad (13)$$

and where

$$\mu_2 = \mu + \frac{\partial \mu}{\partial \epsilon} \delta \epsilon \quad (14)$$

The procedure of Part I may be used with only minor modification to show that the third term becomes

$$\left[\frac{\partial}{\partial x'} \left(\frac{\partial g}{\partial \epsilon} \right) \delta \epsilon \frac{\partial x}{\partial t} + \frac{\partial}{\partial y'} \left(\frac{\partial g}{\partial \epsilon} \right) \delta \epsilon \frac{\partial y}{\partial t} \right]_{s=\zeta}$$

Using equation 11, we may now write equation 12 as

$$\begin{aligned} \frac{d(P_2 - P_1)}{dt} &= \frac{-K}{f_0^2} \left(\frac{1}{q^2} - 1 \right) \\ &\cdot \left\{ \left[N \sqrt{1 + x'^2 + y'^2} \frac{d\zeta}{dt} \right. \right. \\ &+ \left. \frac{N}{\sqrt{1 + x'^2 + y'^2}} \left(x' \frac{\partial x}{\partial t} + y' \frac{\partial y}{\partial t} \right) \right]_{s=\zeta} \\ &+ \int_0^\zeta \frac{\partial N}{\partial t} \sqrt{1 + x'^2 + y'^2} dz \Big\} \\ &+ \left\{ \mu_2 \frac{d\zeta}{dt} \delta(\sec x) \right\}_{s=\zeta} \end{aligned}$$

As in Part I, the term in square brackets may be written in terms of V_{11} , and the integral may be written as an integral along the path. Hence,

$$\begin{aligned} \frac{d(P_2 - P_1)}{dt} &= \frac{K}{f_0^2} \left(1 - \frac{1}{q^2} \right) \left[(N V_{11})_{s=\zeta} \right. \\ &+ \left. \int_\Gamma \frac{\partial N}{\partial t} ds \right] + \mu_2 \delta V_{11} \quad (15) \end{aligned}$$

where

$$\delta V_{11} = \delta \left(\frac{d\zeta}{dt} \sec x \right) = \frac{d\zeta}{dt} \delta(\sec x)$$

Equation 15 is in a form that is independent of the co-ordinate system. The result, however, neglects those terms of equation 7 that involve the differences between the ray paths for the two frequencies, except for the term in δV_{11} .

We can also derive equation 15 from the results of Part I, by making several approximations. Applying equation 13 of Part I to each frequency, and writing $\mu \doteq 1 - (KN/f^2)$, we find

$$\begin{aligned} \frac{d(P_2 - P_1)}{dt} &= [\mu_2 V_{11}^{(2)} - \mu_1 V_{11}^{(1)}] \\ &- \frac{KN}{q^2 f_0^2} V_{11}^{(2)} + \frac{KN}{f_0^2} V_{11}^{(1)} \\ &+ \frac{K}{q^2 f_0^2} \int_\Gamma \frac{\partial N}{\partial t} ds - \frac{K}{f_0^2} \int_\Gamma \frac{\partial N}{\partial t} ds \end{aligned}$$

where the subscripts and superscripts refer to signals 1 and 2. If we now neglect the difference between μ_1 and μ_2 in the first term, we obtain equation 15.

This alternative method for solving the problem has the advantage of greater generality of propagation mode. In deriving equation 15, it was necessary to restrict the form of the refractive index to a form which permitted us to attach a simple significance to the quantity ϵ . Hence, the results are applicable only for quasi-transverse propagation of the ordinary mode. On the other hand, equation 13 of Part I does not depend upon the specific form of the refractive index, and hence the time derivative of $P_2 - P_1$ can be computed for an arbitrary mode of propagation (assuming a quasi-isotropic medium). The terms depending upon the ray slope can be included if necessary.

We note that equation 15 can be solved to yield

$$\begin{aligned} N(z = \zeta) &= \frac{c(f_2 - qf_1)f_0q}{K(q^2 - 1)V_{11}} \\ &- \frac{1}{V_{11}} \int_\Gamma \frac{\partial N}{\partial t} ds - \frac{\mu_2 f_0^2 q^2 \delta V_{11}}{K(q^2 - 1)V_{11}} \quad (16) \end{aligned}$$

In this equation, the first and third terms may be evaluated by using a measurement of the frequency difference $f_2 - qf_1$ and by calculating V_{11} and δV_{11} from the trajectory and an assumed ionosphere. For most purposes $\mu_2 \doteq 1$. The integral term may be considered to be an error term, or, under suitable conditions, may be computed [Kelso, 1960b] from a measurement of the Faraday rotation. As indicated by Kelso, the integral may make a large contribution to equation 16 for a measurement in a vehicle at heights well above the major part of the ionosphere.

4. *Rate of Faraday rotation.* In quasi-longitudinal propagation, the two magneto-ionic modes may be supposed to be circularly polarized with opposite senses of rotation. On entering the ionosphere, a plane-polarized wave may be considered to be decomposed into two such circularly polarized modes of equal amplitudes. As the wave propagates, the two modes, having slightly different refractive indices, travel with different phase velocities. Thus, recombining the two modes at some point along the path, we find that the net polarization is still linear, but that the direction of polarization has rotated about an axis parallel to the direction of propagation, by an angle ψ (in radians) equal to $\frac{1}{2}$ the difference in phase between the two modes. Thus, to determine the rate of Faraday rotation, $d\psi/dt$, we must compute k_0 times the time derivative of the phase-path difference $P_2 - P_1$ of the two modes. To do this we may use the results of section 2. The refractive index for quasi-longitudinal propagation is given by (see *Mitra* 1952] or *Ratcliffe* [1959] for detailed discussion)

$$\mu_{1,2} = 1 - \frac{2KN/f^2}{1 \pm (f_L/f)}$$

where

$$f_L = f_H \cos \theta.$$

$$f_H = (H e \mu_0)/(2\pi m) = \text{gyrofrequency.}$$

H = geomagnetic field intensity in ampere turns/meter.

μ_0 = permeability of free space = 1.257×10^{-8} henry/meter.

θ = angle between direction of ray and geomagnetic field.

The remaining quantities have been defined above.

Assuming high frequencies, μ may be written approximately, as

$$\mu_{1,2} \doteq 1 - \frac{KN}{f^2} \mp \frac{KNf_L}{f^3} \quad (17)$$

where we associate μ_1 with the upper sign, which corresponds to the extraordinary mode.

If we assume the two modes to travel along the same ray path, we may write ψ as

$$\begin{aligned} \psi &= \frac{k_0}{2} (P_2 - P_1) = \frac{k_0}{2} \int_{\Gamma} (\mu_2 - \mu_1) ds \\ &= \frac{Kk_0}{f^3} \int_{\Gamma} Nf_L ds \quad (18) \end{aligned}$$

Then, $d\psi/dt$ may be found by differentiating (18) in the manner illustrated in Part I. Instead, we shall apply the results of section 2.

Write

$$\mu = 1 - \frac{KN}{f^2} + \epsilon \left(\frac{KNf_L}{f^3} \right)$$

In going from μ_1 to μ_2 , $\delta\epsilon = 2$, and we have

$$\frac{d\mu}{d\epsilon} = \frac{KNf_L}{f^3} \quad (19)$$

We must note here that f_L is a function of the slope of the ray, and is, therefore, a function of x' and y' . Substituting equation 19 into equation 8,

$$\begin{aligned} \frac{d\psi}{dt} &= \frac{k_0}{2} \frac{d(P_2 - P_1)}{dt} \\ &= \frac{2\pi K}{cf^2} \left(Nf_L \frac{d\zeta}{dt} \sqrt{1 + x'^2 + y'^2} \right)_{s=\Gamma} \\ &\quad + \int_0^{\Gamma} \left\{ \left[\frac{\partial}{\partial x} (Nf_L) \frac{\partial x}{\partial t} \right. \right. \\ &\quad \left. \left. + \frac{\partial}{\partial y} (Nf_L) \frac{\partial y}{\partial t} \right] \sqrt{1 + x'^2 + y'^2} \right. \\ &\quad \left. + N \left[\frac{\partial}{\partial x'} (f_L \sqrt{1 + x'^2 + y'^2}) \frac{\partial x'}{\partial t} \right. \right. \\ &\quad \left. \left. + \frac{\partial}{\partial y'} (f_L \sqrt{1 + x'^2 + y'^2}) \frac{\partial y'}{\partial t} \right] \right. \\ &\quad \left. + \frac{\partial}{\partial t} (Nf_L \sqrt{1 + x'^2 + y'^2}) \right\} dz \\ &\quad + \left[\mu_2 \frac{d\zeta}{dt} \delta \sec x \right]_{s=\Gamma} \quad (20) \end{aligned}$$

We may now combine terms and apply Euler's equation as we did in the previous case of the two-frequency Doppler experiment. Recalling that f_L is a function of x' and y' , but assuming that it is not an explicit function of t , we obtain,

$$\begin{aligned} \frac{d\psi}{dt} &= \frac{k_0}{2} \frac{d(P_2 - P_1)}{dt} \\ &= \frac{2\pi K}{cf^2} \left[(Nf_L V_{11})_{s=\Gamma} + \int_{\Gamma} f_L \frac{\partial N}{\partial t} ds \right. \\ &\quad \left. + \int_{\Gamma} N \left(\frac{\partial f_L}{\partial x'} \frac{\partial x'}{\partial t} + \frac{\partial f_L}{\partial y'} \frac{\partial y'}{\partial t} \right) ds \right] \\ &\quad + \frac{k_0}{2} \mu_2 \delta V_{11} \quad (21) \end{aligned}$$

Since we have used equation 8 instead of 7, the results in equations 20 and 21 do not include the effects caused by the fact that the two modes follow different paths except as introduced by the last term. For any specific example, we can use equation 7 to test for the magnitude of the errors made by the neglect of the difference in paths. In general, this effect will be small for reasonably high frequencies.

Equation 21 is a more precisely stated version of the corresponding result as given by Kelso, [1960b]. In the earlier statement, the two integrals of equation 21 were represented as a partial derivative with respect to time of the integral of Nf_L along the path. This loose notation, as stated in the original reference, was used to indicate the effects of changes in the ionosphere (explicit functions of time, corresponding to the first integral of equation 21, and the changes caused by motions across the ray path, corresponding to the second integral of equation 21).

When the ray is nearly parallel to the magnetic field, the quantities $\partial f_L / \partial x'$ and $\partial f_L / \partial y'$ are both small. Also, for vehicle motions largely in the ray direction, the quantities $\partial x' / \partial t$ and $\partial y' / \partial t$ are small. Under these circumstances, the second integral may be neglected, and equation 21 will reduce to the first two terms on the right-hand side. Using a mean value \bar{f}_L for f_L , in the integral term, we have

$$\frac{d\psi}{dt} = \frac{2\pi K}{cf^2} \left[(Nf_L V_{11})_{z=\zeta} + \bar{f}_L \int_{\Gamma} \frac{\partial N}{\partial t} ds \right] + \frac{k_0}{2} \mu_2 \delta V_{11} \quad (22)$$

Except for the presence of f_L and \bar{f}_L , the factor in brackets [] is the same as the corresponding factor in equation 15. Evaluating f_L (at $z = \zeta$) and \bar{f}_L , equations 15 and 22 can be solved simultaneously for N ($z = \zeta$) and the integral $\int (\partial N / \partial t) ds$ when the paths are in a form so that δV_{11} is negligible. Therefore, as suggested by Kelso, a combination of a two-frequency Doppler experiment with a measurement of the rate of Faraday rotation can be used to eliminate the error caused by $\partial N / \partial t$. The present description of the treatment of the combined experiment is very sketchy. An actual application would require careful attention to estimation of the errors caused by the neglect of the discarded

terms, and the error caused by the estimation of \bar{f}_L .

It is of interest to use our results to find $d\psi/dt$ for a special form of ionosphere. Let us assume that N is a function of z alone, that the rays are straight, and that f_L is independent of height. These assumptions are valid for high enough frequencies, and when the ray is not too far from the vertical.

Then, for a vehicle at ξ, η, ζ we assume the motion to be in the xz plane, and write the velocity \mathbf{v} as

$$\mathbf{v} = i v_x + k v_z = i \frac{d\xi}{dt} + k \frac{d\zeta}{dt}$$

the radius vector to the vehicle is

$$\mathbf{r} = \frac{i\xi + j\eta + k\zeta}{\sqrt{\xi^2 + \eta^2 + \zeta^2}}$$

Let the magnetic field be given by

$$\mathbf{H} = iH_x + jH_y + kH_z$$

If χ is the angle between the ray and the vertical (z direction), we have

$$\sec \chi = \frac{\sqrt{\xi^2 + \eta^2 + \zeta^2}}{\zeta}$$

Clearly

$$x' = \frac{\xi}{\zeta}$$

$$y' = \frac{\eta}{\zeta}$$

Then

$$\begin{aligned} f_L &= \left(\frac{e^2 \mu_0}{2\pi m} \right) \mathbf{H} \cdot \mathbf{r} = \frac{e \mu_0}{2\pi m} \left(\frac{\xi H_x + \eta H_y + \zeta H_z}{\sqrt{\xi^2 + \eta^2 + \zeta^2}} \right) \\ &= \frac{e \mu_0}{2\pi m} \left(\frac{x' H_x + y' H_y + H_z}{\sqrt{x'^2 + y'^2 + 1}} \right) \end{aligned}$$

and

$$\begin{aligned} f_L \sqrt{1 + x'^2 + y'^2} \\ = \frac{e \mu_0}{2\pi m} (x' H_x + y' H_y + H_z) \end{aligned}$$

Let $\partial N / \partial t = 0$. Substituting into equation 22 we find

$$\frac{d\psi}{dt} = \frac{K}{f^3} \frac{\epsilon\mu_0}{2\pi m} \left[\left(N H_L \sec \chi \frac{dz}{dt} \right)_{z=\xi} + \left(H_x \frac{\partial x'}{\partial t} + H_y \frac{\partial y'}{\partial t} \right) \int_0^\xi N dz \right]$$

it,

$$\frac{\partial x'}{\partial t} = \frac{v_x}{\xi} - \frac{\xi v_z}{\xi^2} \quad \frac{\partial y'}{\partial t} = -\eta v_z / \xi^2$$

thus,

$$\frac{d\psi}{dt} = \frac{K\epsilon\mu_0}{mf^2c} \left[\left(N H_L \sec \chi \frac{dz}{dt} \right)_{z=\xi} - \frac{H_x v_x}{\xi} \int_0^\xi N dz - \frac{(\xi H_x + \eta H_y) v_z}{\xi^2} \int_0^\xi N dz \right]$$

This result agrees with that obtained by Garriott, [1959] for this case. If, also, $v_z = 0$, the result reduces to an expression obtained earlier [Bowhill, 1958] for horizontal motion.

In obtaining this result, we used equation 20 instead of 21. These two equations will not give exactly the same expression for this simple example, since the use of a straight ray for a problem in which N varies with height is only an approximation. In going from (20) to (21) we used Euler's equations, which apply to the true path. One would not expect to obtain the same expressions by applying a straight ray approximation before and after such a transformation. On the other hand, for a medium for which N is constant, the rays are truly straight, and both (20) and (21) lead to the result that

$$\frac{d\psi}{dt} = \left(\frac{K\epsilon\mu_0}{mcf^2} \right) N (H_{x,r} + H_{y,r})$$

Thus, we see that the present results reduce in special cases to the forms obtained by other authors.

Acknowledgments. I wish to express my indebtedness to Dr. Allen R. Sims and Dr. Carl D. Graves, of Space Technology Laboratories, and to Dr. O. K. Garriott, of Stanford University, for numerous and valuable discussions with them concerning various aspects of the present work.

REFERENCES

- Bowhill, S. A., *J. Atmospheric and Terrestrial Phys.*, **13**, 175-176, 1958.
 Browne, I. C., J. V. Evans, J. K. Hargreaves, and W. A. S. Murray, *Proc. Phys. Soc., B*, **69**, 901-920, 1956.
 Daniels, Fred B., and Siegfried J. Bauer, *J. Franklin Inst.*, **267**, 187-200, 1959.
 Garriott, Owen K., *Technical Report No. 1*, Radio Propagation Laboratory, Stanford University, 1959.
 Jackson, John E., and J. Carl Seddon, *J. Geophys. Research*, **63**, 197-208, 1958.
 Kelso, John M., *J. Geophys. Research*, **65**, 3909-3914, 1960a.
 Kelso, John M., *Electromagnetic Wave Propagation*, edited by M. Desirant and J. L. Michiels, Academic Press, London and New York, 1960b.
 Mitra, S. K., *The Upper Atmosphere*, 2d ed., The Asiatic Society, Calcutta, 1952.
 Ratcliffe, J. A., *The Magneto-Ionic Theory*, Cambridge University Press, Cambridge, 1959.
 Seddon, J. Carl, *J. Geophys. Research*, **58**, 323-335, 1953.

(Manuscript received November 14, 1960;
 revised January 18, 1961.)

Ionization Loss Rates Below 90 Km

CULLEN M. CRAIN

*The RAND Corporation
Santa Monica, California*

Abstract. Quasi-equilibrium solutions for electron and ion density below 90 km are examined. Special attention is given to the present uncertainty in some of the rate coefficients and to the differences between daytime and nighttime results.

INTRODUCTION

The purpose of this paper is to survey the current state of understanding of certain ionization phenomena in the normal atmosphere from ground level to about 90 kilometers. Primarily, the paper deals with the various electron and ion (positive and negative) decay rates and the concentration of these charges resulting from given electron production rates in the above altitude range. Simple, approximate expressions are derived for day and night quasi-equilibrium conditions and comparisons given for the relative effects of ionizing radiation during daytime and nighttime.

IONIZATION RATE EQUATIONS

Consider first the widely used differential equations for rate of change of charge density in the atmosphere

$$\frac{dN}{dt} = P - AN - BN(N + N^-) + CN^- + DN^- \quad (1)$$

$$\frac{dN^-}{dt} = AN - (C + D)N^- - EN^-(N + N^-) \quad (2)$$

$$\frac{dN^+}{dt} = \frac{dN}{dt} + \frac{dN^-}{dt} \quad (3)$$

Where P is the electron production rate, N^+ , N^- , and N are the positive ion, negative ion, and electron density, respectively,

A is the electron attachment coefficient.

B is the electron-ion recombination coefficient.

C is the photo-detachment coefficient.

D is the collisional detachment coefficient.

E is the ion-ion recombination coefficient.

The coefficients A , B , C , D , and E as used above are naturally functions of altitude. Their values as a function of altitude will be considered later. Adding equations 1 and 2 gives

$$\frac{dN^+}{dt} = P - BN(N + N^-) - EN^-(N + N^-) \quad (4)$$

For situations where the rate of change of charge density is slow compared to the various rates on the right-hand side of equations 1, 2, and 4 the left-hand side of the equations can reasonably be set equal to zero. Thus, for the near equilibrium condition equations 1, 2, and 4 become respectively

$$P = AN + B(N + N^-)N - (C + D)N^- \quad (5a)$$

$$0 = AN - (C + D)N^- - EN^-(N + N^-) \quad (5b)$$

$$P = BN(N + N^-) + EN^-(N + N^-) \quad (5c)$$

Dividing equation 5b by N ,

$$A = (C + D)N^-/N + E(N^-/N)(N + N^-) \quad (6)$$

Defining N^-/N as λ , and rewriting (6),

$$\lambda = A/(C + D + EN^+) \quad (7)$$

Equation 7 obviously applies only for situations where the assumption that dN^-/dt is much less than AN or $(C + D)N^- + EN^-(N + N^-)$ is valid. For short intervals after gross perturbations in the atmospheric ionization density this assumption will naturally be invalid and the complete rate equations must be used.

RATE COEFFICIENTS

(A) *Electron attachment.* Of the principal gaseous constituents of the normal atmosphere up to about 90 km only oxygen and water vapor form negative ions by electron attachment to the molecular species. Rarer species such as ozone, NO, N₂O, NO₂, O, etc., also are known to attach electrons; however, below 90 km in the normal atmosphere there is no evidence to indicate that these species can compete with the more abundant oxygen and water vapor molecules in determining the effective attachment rate. Since water vapor even near the earth's surface never exceeds about 1/5 of the oxygen concentration, and since its concentration decreases rapidly with altitude, it would be necessary that water vapor have a considerably higher attachment cross section than oxygen for water vapor to be effective in controlling the attachment rate. While no pertinent data (i.e., at thermal energy) are available for water vapor, this possibility is considered unlikely; however, it should be examined experimentally. In view of the above discussion, the best approach at present for estimating the coefficient A at any altitude up to 90 km appears to be that of considering the effect of O₂ alone. Unfortunately the picture is not clear in this regard. While attachment rates for various pressures and electron energies have been investigated experimentally for several decades, no data were obtained for electrons with thermal energy before microwave measurements of *Biondi* [1951]. *Biondi* interpreted his data as indicating an attachment cross section of 1.2×10^{-22} cm² in the pressure range 8 to 25 mm of mercury. Measurements by *Crain* [1958] of the effects of radiation from a nuclear reactor and from a Co⁶⁰ source on the resonant frequency of a 3-cm cavity resonator containing air and other gases indicated the value reported by *Biondi* was orders of magnitude too small in the pressure range from 1 to 10^{-4} atmospheres; however, these measurements were probably diffusion limited below about 10^{-3} atmospheres. More recent measurements by *Chanin, Phelps, and Biondi* [1959], using the drift-tube technique at low electron energies have indicated that for thermal electrons at 300°K, and for pressures between approximately 5 and 50 mm of Hg, the attachment frequency for atmospheric air can be expressed as follows:

$$A = 3000 P_{O_2} p^2 / \text{sec} \quad (8)$$

where P_{O_2} is the oxygen pressure in millimeters of mercury. Expressed in terms of atmospheric pressure in atmospheres, the above becomes

$$A = 7 \times 10^7 p^2 / \text{sec} \quad (9)$$

Recent measurements at the University of Texas (*Deam*, private communication) using a 400-Mc refractometer, a Co⁶⁰ source and the technique of *Crain* [1958] have given lower electron densities than an A value of $7 \times 10^7 p^2$ would predict.

These most recent measurements at about 0.01 atmosphere of air can be explained satisfactorily if $A \cong 2 \times 10^8 p^2$ for $p = 0.01$ atmosphere. Such a value would also be compatible with the results of *Crain* [1958]. Measurements on atmospheric air at the General Electric Company [*Fitzpatrick*, 1959] using a cavity resonator and a pulsed high-energy electron beam have indicated a value of A that is proportional, in an air-pressure range of 5 to 60 mm. to pressure to about the 4/3 power. Although this pressure dependence is not predicted by current theory, it is most interesting to note that at an air pressure of 0.01 atmosphere the General Electric results give a value of $A = 2 \times 10^8$ (about the same as the University of Texas results at the same pressure) and at an air pressure of 0.1 atmosphere, they gave an A value of 4.5×10^8 or somewhat less than those deduced by *Chanin, Phelps, and Biondi* [1959] for the same oxygen partial pressure. It appears that all recent measurements using drift-tube steady-state radiation of cavity resonators, and pulsed high-energy electron beam technique agree, for an air pressure range of 1.0 to 0.0 atmospheres, to within a factor of about 3 to 10 following:

$$A = 10^8 p^2 \quad (9c)$$

An urgent need for experimental investigation of the range of pressure for which the formula may be used remains; however, order of magnitude errors would not be expected for $p > 10^{-4}$. At pressures less than about 10^{-4} or 10^{-5} two-body type attachment (radiative) probably becomes important. Unfortunately, the value for radiative attachment is not available from direct measurements. Cross sections from about 10^{-19} to 10^{-26} /cm² have been suggested as possibilities in the literature. If one arbitrary

assumes a value of $2 \times 10^{-24}/\text{cm}^2$, the resulting value of A is as follows:

$$A = 10^8 p^2 + 10^2 p \quad (10)$$

Obviously, general application of the above expression to the atmosphere is subject to question, particularly for pressures less than 10^{-3} atmospheres, and for secondary electrons that are produced by various ionizing radiations and hence pass through a wide range of energy while thermalizing. It is doubtful, however, if evidence available currently from laboratory or ionospheric studies can justify an appreciably (order of magnitude) different expression for altitudes below about 90 km.

B. Electron-ion recombination. Laboratory measurements have suggested that the recombination of electrons with ions of diatomic molecules such as N_2^+ and O_2^+ occurs at pressures of a few millimeters or less primarily through a dissociative process such as



The coefficients for such reactions have been reported by *Biondi and Brown* [1949] as 2.8×10^{-7} for O_2 and 1.4×10^{-6} for N_2 . Other investigators have reported values for N_2 which range from 10^{-6} to 4×10^{-7} .

It is apparently the consensus of opinion that the recombination rate in the normal E region should be controlled by dissociative recombination of O_2^+ and electrons. Until recently the generally used value for E region recombination has been near 10^{-8} . Hence, there has been some difficulty of reconciling laboratory observations with those deduced from ionospheric measurements [*Gerjuoy and Biondi*, 1953]. Recent X-ray flux measurements by *Friedman* and others seem to clarify the situation. The value of 10^{-8} has been arrived at primarily from observing ionospheric behavior during solar eclipses and by assuming that at totality the source function is zero. *Friedman's* rocket observations during eclipses have shown that his assumption is quite erroneous. The net result is that recombination coefficients derived from eclipse observations naturally are too low if the source function is assumed to become zero when in reality it only decreases to a smaller value. *Friedman* [1959] reported that a value of 10^{-7} for the E region now appears more appropriate. *Piddington* [1951] argued that for several

reasons a 10^{-7} value may be more realistic than 10^{-8} . It would appear then that laboratory and ionospheric observations are now reasonably compatible and that a dissociative recombination value for O_2^+ of about 10^{-7} is satisfactory. This value should also be normally appropriate for the region below 100 km since the three-body recombination coefficient for air is reported to be only about 1.7×10^{-7} at STP. For ionization of air by sources which produce appreciable N_2^+ ions the value of 10^{-7} is perhaps too low since the N_2^+ dissociative recombination coefficient is apparently somewhat greater ($4 \times 10^{-7}/\text{sec}$) [*Faire and Champion*, 1959]. It would appear, however, that N_2^+ ions will rapidly charge exchange with O_2 molecules to produce O_2^+ ; hence, the resulting recombination coefficient is in general difficult to predict precisely. A value of 10^{-7} appears at present the most reasonable value for the electron-ion recombination coefficient below 100 km.

C. Photo-detachment. The coefficient for photo-detachment of electrons from O_2 ions is given [*Mitra*, 1952]:

$$C = \int_{\nu_0}^{\infty} Q(\nu) N_{\nu} d\nu \quad (12)$$

where

$Q(\nu)$ = cross section for light absorption of frequency ν .

ν_0 = light frequency for the minimum energy of detachment.

N_{ν} = number of light quanta of frequency ν passing through unit area per unit time.

Attempts to measure $Q(\nu)$ by different means in the last few years have led to fairly widespread experimental values. When the various results are interpreted in terms of electron affinity to the O_2^- ions, values have ranged from about 0.15 ev to 0.9 ev. Recent measurements of *Smith, Burch, and Branscomb* [1958] have indicated the lowest affinity in the above range. From their measurements it is possible to calculate a value of C for solar radiation of about 0.4 per sec; however, there are some questions about the validity of applying their experimental values to the ionized atmosphere. These questions center around whether the O_2^- ions for which their laboratory values were obtained were in the same electronic state as O_2^- in the normal

atmosphere. At present, however, the value of about 0.4 per sec appears the best available for sunlight.

D. Collisional detachment. Since below about 80 km the principal negative ion in the atmosphere is O_2^- , any significant electron contribution via collisional detachment would be expected to arise from this ion. Neither theory nor current data are yet satisfactory for confidently specifying the rate at which electrons may be detached from O_2^- in the atmosphere. Uncertainties in the electron affinity of O_2^- reflect very heavily into the problem. Supposedly, if we knew the electron affinity of the normal atmospheric O_2^- ion, we could calculate with useful accuracy a detachment rate coefficient appropriate for collision with the various other molecular species involved. In general the detachment rate should be proportional to $e^{-(V_a/kT)}$. If the affinity were, say, 0.15 ev and $kT \cong 1/40$ ev, then

$$D = K e^{-15/40} = K e^{-0.375} = K/400.$$

However, for an affinity of 0.90 ev,

$$D = K e^{-9 \times 40} = K e^{-36} \cong K/(8,000)^4 \cong K/(4 \times 10^{15}).$$

Obviously, calculating detachment rates from the electron affinity requires considerable accuracy in the affinity value if the deduced detachment rate is to be meaningful. Each 0.1 ev of error in the affinity introduces an error factor of about 55 in the detachment rate (at a temperature of 300°K).

Recent attempts have been made to deduce collisional detachment rate coefficients from ionospheric absorption observations following solar flares and from various laboratory measurements. Branscomb and Bailey deduced a value of 0.25 ev for electron affinity from such ionospheric absorption measurements. They further suggested that a rate coefficient D of about $2 \times 10^{-17}n$, where n was the number of air molecules per cubic centimeter, was appropriate. Biondi and Phelps [1959] have attempted to deduce detachment values from their drift-tube swarm measurements and have preliminarily reported that the results show 'that the frequency of detachment collisions for the negative ions in our experiment is less than 3 p sec⁻¹ at 300°K, where p is the oxygen pressure in mm of mercury. If this result is combined with the

measured attachment frequency of 300°K, i.e., 2500 p^2 , and substituted into the law of mass action, one finds that the affinity of the negative ion formed in our experiment is greater than 0.3 electron volt. This value is significantly larger than the value of 0.15 ev obtained by extrapolation to threshold of the photo-detachment measurements of Smith, Burch, and Branscomb, or the value of less than 0.25 ev required to give the collisional detachment frequency postulated by Bailey and Branscomb to explain certain features of the absorption of radiowaves in the D layer.' The previously discussed values of attachment rate deduced from the University of Texas 400-Mc refractometer measurements were based on the assumption of negligible detachment. If detachment were contributing to the result, a higher rate of attachment must be assumed than was given in section A. We can deduce approximately from the University of Texas measurements the level of collisional detachment that would have been significant in the evaluation of the attachment coefficient A . Reference to equation (1) shows that for no photo-detachment (the metallic cavity was completely enclosed) $\lambda = A/(D + EN^+)$. At about 0.01 atmosphere, and at the radiation levels used in the experiment EN^+ for oxygen was near 10. The value of D that could then be considered significant at 0.01 atmosphere of O_2 would be approximately 1. A value for D of 1 for $p = 7.6$ mm would mean, using Biondi's above nomenclature, a detachment frequency of 0.13 P_{O_2} , where P_{O_2} is the oxygen pressure in millimeters of Hg. The corresponding electron affinity would be greater than 0.38 ev. Certainly a value as low as 0.3 ev would have greatly affected the measurement, and, as stated before, would require a much higher attachment rate than deduced in section A. If, then, use is made of an A value as given earlier ($10^8 p^2 + 10^2 p$), the above argument would require a D value no greater than about 100 p , where p is in atmospheres. Actually one can argue that a more reasonable value should be much lower since the evidence for electron affinities >0.38 ev for O_2^- cannot, as yet, be discarded. It turns out that in many practical problems one can just as well take $D \cong 0$: $D = 100 p$ since even this value is often insignificant compared to $(EN^+ + C)$, C , or EN^+ . The above material is the basis for the later co-

usions that at nighttime the rate equations can be solved, assuming negligible detachment, and that the results given in Figures 3 and 5 are appropriate. The writer has been unable to find experimental evidence that conflicts with the practical application of the value $D \cong 0$ in the ionization rate equations for the normal atmosphere. Clearly, however, the true value of this parameter is presently not satisfactorily solved as a function of atmospheric density or temperature.

E. Ion-ion recombination. In 1924 J. J. Thompson published a theory for ion-ion recombination by a three-body process. This theory has been found quite in agreement with experiment in the pressure range from about an atmosphere to a few millimeters and is reviewed by Loeb [1955, pp. 545-550]. Above about 20,000 ft Thompson's three-body coefficient can be reduced

$$E = 6p/T^{5/2} \quad (13)$$

where p is the pressure in atmospheres and T is the temperature in degrees K.

At low pressures the above expression becomes vanishingly small (about 10^{-10} at 80 km) and other mechanisms are thought to become more effective in controlling ion-ion recombination. Loeb [1955, p. 558] indicates Thompson's expression should apply down to about 10^{-2} mm and that at 10^{-3} mm the value probably levels off

at 10^{-10} . He admits that a value as high as 3×10^{-9} is possible. Massey and Burhop [1956, p. 622] have calculated that the contribution to recombination from charge neutralization of O^+ and O^- could be between 5×10^{-8} and 5×10^{-9} and that a larger value is unlikely. They do not appear to have made similar calculations for O_2^+ and O_2^- . It is not clear to the writer at what altitude the above values are appropriate since the O concentration is less than one part per thousand even at 80 km. Massey and Burhop state that 'no experimental evidence is available to test their theoretical conclusions and it is difficult to see how any direct evidence could be obtained.' Loeb indicates the 'urgent need for this sort of study as it bears on problems encountered in the upper atmosphere' but indicates the problem in 'very difficult.'

In view of the above, it appears reasonable currently to adopt for E the following expression based on Thompson's three-body value plus a two-body term in the range suggested by Loeb:

$$E = 6p/T^{5/2} + 10^{-9} \quad (14)$$

where p is in atmospheres.

QUASI-EQUILIBRIUM DENSITIES

The various rate coefficients discussed above are plotted as a function of altitude in Figure 1 for altitudes between 50 and 90 km. The value

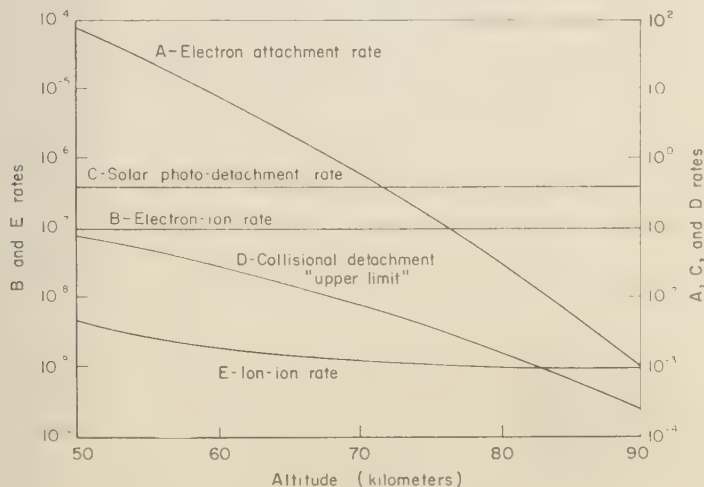


Fig. 1. Rate coefficients vs. altitude.

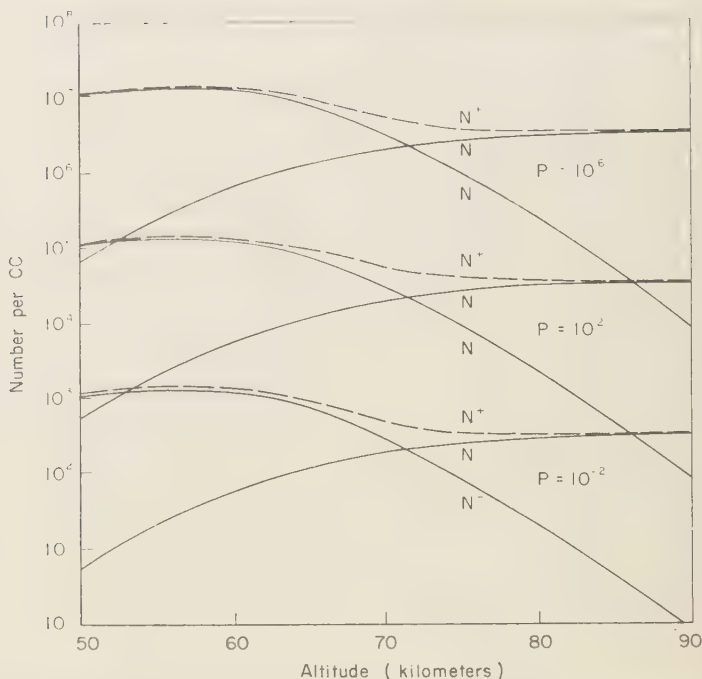


Fig. 2. Daytime equilibrium electron, positive ion, and negative ion densities vs. altitude for various electron production rates.

plotted for collisional detachment D is the 'upper limit' one discussed previously.

The quasi-equilibrium charge densities at any altitude and for any steady-state production rate can be calculated directly by substituting the appropriate values of the rate coefficients of Figure 1 into equation 5. Three particular cases of interest are shown in Figures 2, 3, and 4. In Figure 2 the value of photo-detachment appropriate for daylight conditions has been used. For this case it is seen from Figure 1 that in the altitude range 50- to 90-km photo-detachment is always appreciably greater than the 'upper limit' of collisional detachment; hence, collisional detachment has negligible effect. For Figure 3 both photo-detachment and collisional detachment were taken as negligible; hence, the electron densities shown in Figure 3 represent the lowest electron density values expected for various nighttime production rates. For Figure 4 the photo-detachment rate was taken as zero and the 'upper limit' of collisional detachment, i.e., $D = 100 p$, where p is density in atmospheres,

was used; hence, the electron densities of Figure 4 represent the maximum nighttime electron densities to be expected for the various production rates shown.

In Figure 5 the electron density results in Figures 2 and 3 are compared to show, for the various production rates, the maximum ratio of day-to-night electron density. As is obvious in the figure, the ratio is a function of both altitude and the production rate. At low production rates very high day-to-night ratios are to be expected if collisional detachment is negligible.

Figure 6 compares the electron densities of Figures 2 and 4, and thus gives the *minimum* day-to-night electron density ratios, for a range of electron production rates. Available experimental evidence indicates that for production rates in the vicinity of 10^2 electrons/cm²/sec the day-to-night ratio must be larger than the values indicated by Figure 6 and actually near those deduced in Figure 5; hence, we must currently conclude that the negligible collisional detachment result is the most satisfactory.

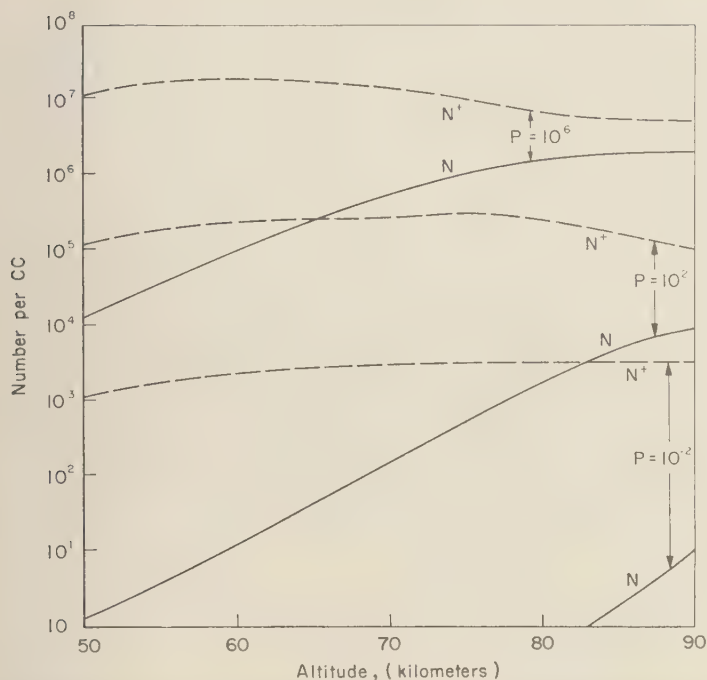


Fig. 3. Equilibrium electron and positive ion densities vs. altitude for various electron production rates and for negligible detachment.

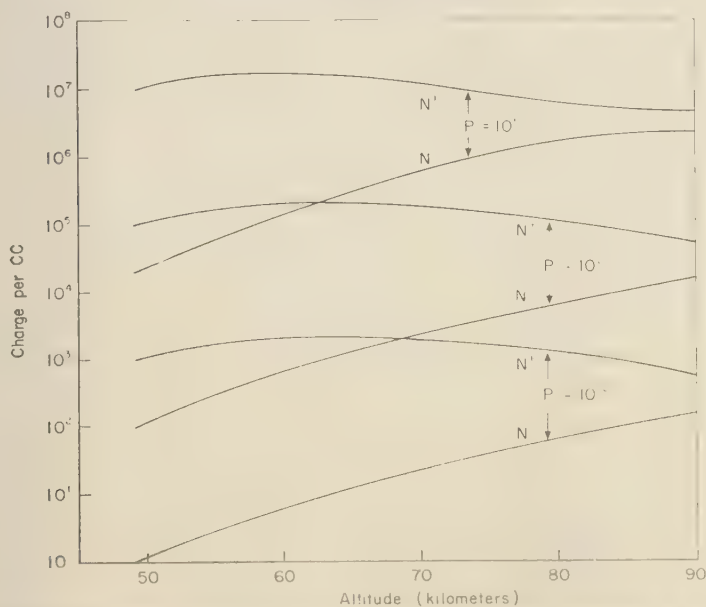


Fig. 4. Equilibrium electron and positive ion densities for collisional detachment $\approx 100 p$ and negligible photo-detachment.

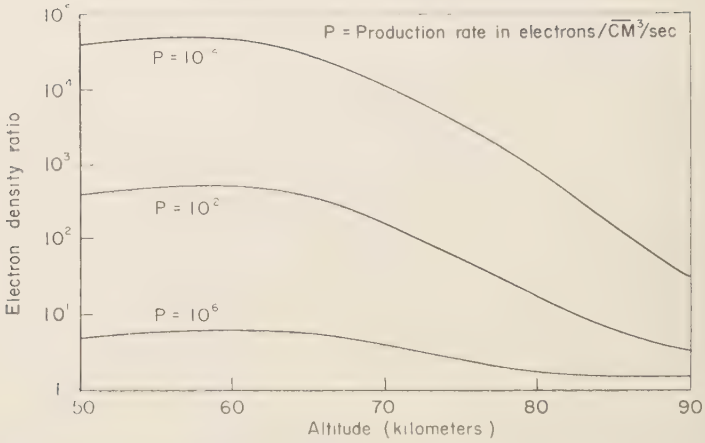


Fig. 5. Ratio of electron densities for photo-detachment and negligible detachment cases (maximum day-to-night electron density rates for the production rates shown).

EFFECTIVE LOSS RATE COEFFICIENTS

It is common practice in papers dealing with the ionosphere to specify an effective loss rate coefficient for a given region or altitude. Normally this coefficient is taken as follows

$$R_{eff} = P/N^2 \quad (15)$$

where P is the electron production rate, and N , the resulting equilibrium electron density.

While use of such a simple relation is expedient,

it also can be misleading or erroneous; we will pursue this point next.

In equation 7 the exact equilibrium ratio of negative ion to electron density is seen to depend not only on various attachment and detachment parameters but also on the positive ion density. Only where EN^+ is negligible compared to $C + D$ does λ become a so-called 'characteristic of the atmosphere.' In the daytime and for production rates up to and well above those which normally exist between 50 and 90 km the photo

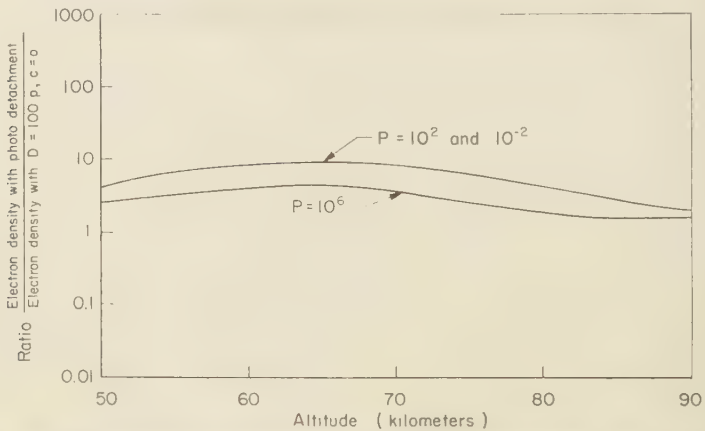


Fig. 6. Ratio of electron densities for cases of solar photo-detachment and for no photo-detachment with 'upper limit' of collisional detachment (minimum day-to-night electron density ratio).

detachment term C dominates the denominator and λ can be taken as simply $\lambda = A/C = N^-/N$.

Let us now return to equation 5. For daytime in the lower ionosphere the dominant terms in (5a) and (5b) are those involving attachment A and detachment C . These equations are thus controlled by the approximate balancing of these terms or

$$O = AN - CN^- \quad (16)$$

To relate the small terms correctly it is necessary to eliminate the large terms (AN and CN^-) from the equations before applying approximations. This may be done by adding (5a) and (5b), thereby obtaining

$$P = BN(N^- + N^+) + EN^-(N + N^+) \quad (5c)$$

This is an exact equilibrium expression relating the small terms, and in this equation one may substitute the appropriate value for N^- , i.e., $N^- = \lambda N = A/C$, and thus obtain

$$\begin{aligned} P &= BN^2(1 + \lambda) + EN^2(1 + \lambda) \\ &= (1 + \lambda)(B + \lambda E)N^2 \\ \text{or} \\ P &= (1 + A/C)(B + A/CE)N^2 \\ &= R_{eff} N^2 \end{aligned} \quad (17)$$

The value of R_{eff} is readily evaluated at a particular altitude using the values of Figure 1.

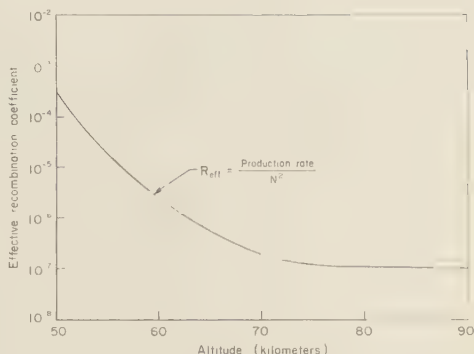


Fig. 7. Daytime loss rate coefficient 50 to 90 km.

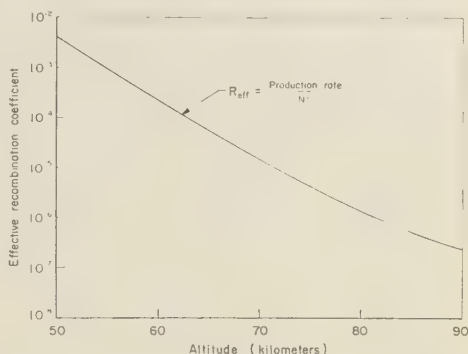


Fig. 8. Minimum expected nighttime loss rate coefficient.

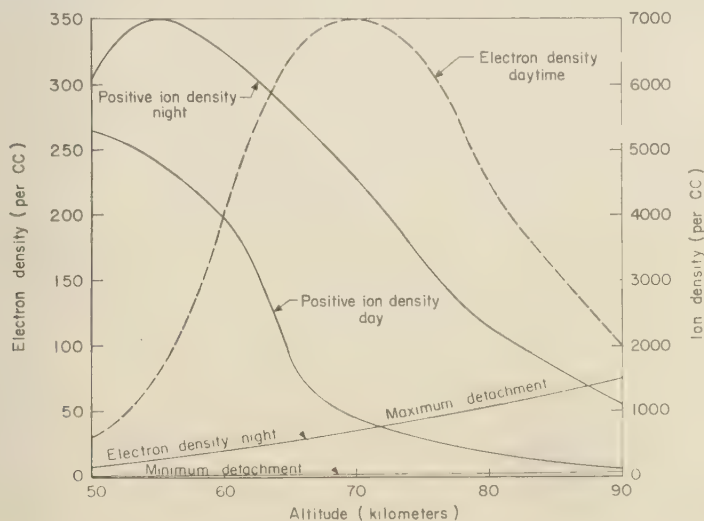


Fig. 9. Electron and positive ion density due to cosmic rays (typical 55°).

Figure 7 plots the resulting value for daytime conditions and for production rates such that EN^+ is negligible in equation 7.

Effective loss rate coefficient in the sense indicated by equation 16 has meaning only if the detachment terms in (7) exceed the term EN^+ . For the values of the parameters adopted in Figure 1 it turns out that if, indeed, D were as high as $100 p$, then, even at night, detachment would dominate the denominator of equation 7 for production rates up to about 10^3 electrons/cc/sec for altitudes below 90 km, and the same procedure as used above for calculating R_{eff} in the daytime would likewise be appropriate at night. The value of R_{eff} has been calculated on this basis and is shown in Figure 8. The electron densities that are calculated using the R_{eff} of Figure 8 are to be considered an upper limit possible at night; also the production rate must be less than about 10^3 at 90 km and less than 10^2 at 65 km for the result to be valid.

Actually the author is convinced that the more accurate electron density value at night is that obtained using negligible detachment. Unfortunately no 'characteristic parameter' such as recombination coefficient or loss rate can be defined and calculated; however, for the range of production rates encountered in the normal atmosphere the attachment term in equation 5a dominates the electron ion recombination terms and the equilibrium electron density is essentially given by the relation

$$N = \frac{\text{Production Rate}}{\text{Attachment Rate}} = \frac{P}{A} \quad (18)$$

For very high production rates one must resort to solution of equation 5 such as was done in plotting the $P = 10^6$ curve in Figure 3.

COSMIC-RAY EFFECTS IN THE LOWER D REGION

As an example of the application of the results of this paper we will consider how the D -region ionization due to cosmic rays alone varies between day and night. For simplicity we will assume that the electron production rate is proportional to atmospheric density in the altitude range 50 to 90 km, and use average cosmic-ray flux values appropriate to a geomagnetic latitude of 55° . Available data indicate

an electron production rate of some 300 electrons/cc/sec/atm is appropriate for this latitude [*Handbook of Geophysics*, 1957].

The electron density has been calculated and plotted in Figure 9 for three conditions as follows: (1) daytime; (2) nighttime assuming negligible collisional detachment; (3) nighttime assuming the 'upper limit' of collisional detachment.

Also shown in Figure 9 are the positive ion densities for daytime and nighttime (negligible detachment solution).

The results in Figure 9 show clearly the large change in electron density expected at altitudes near 70 km due to the photo-detachment effects of solar radiation and are independent of possible additional ionization due to ionizing radiation from the sun. The qualitative effects on VLF radio wave propagation of this phenomena have been previously discussed by Chapman and Davies [1958].

REFERENCES

- Biondi, M. A., *Phys. Rev.*, **84**, 1072, 1951.
- Biondi, M. A., and S. C. Brown, *Phys. Rev.*, **76**, 1697-1700, 1949.
- Biondi, M. A., and A. V. Phelps, *Technical Program Report No. 15*, Physics Dept., Westinghouse Research Laboratories, September 1959.
- Chanin, L. M., A. V. Phelps, and M. A. Biondi, *Phys. Rev. Letters*, **2** (8), 344-346, 1959.
- Chapman, C., and K. Davies, *J. Atmospheric and Terrestrial Physics*, **13**, 86-89, 1958.
- Crain, C. M. J. *Appl. Phys.*, **29**, 1605, 1958.
- Faire, A., and K. Champion, *Phys. Rev.*, **113**, 1-6, 1959.
- Fitzpatrick, J. P., *General Electric Report No. ARDC-TN-59-1*, January 1959.
- Friedman, H., Paper presented at URSI Spring Meeting, Washington, D. C., May 1959.
- Gerjuoy, E., and M. A. Biondi, Dissociative recombination in the E -layer, *J. Geophys. Research*, **58**, 295-303, 1953.
- Handbook of Geophysics*, AFCRC, 1957.
- Loeb L., *Basic Processes of Gaseous Electronics*, University of California Press, 1955.
- Massey, H. S. W., and E. H. S. Burhop, *Electronic and Ionic Impact Phenomena*, Oxford University Press, 1956.
- Mitra, A. P., *The Upper Atmosphere*, The Asiatic Society, Calcutta, 1952.
- Piddington, J. H., The modes of formation of the ionospheric layers, *J. Geophys. Research*, **56**, 409, 1951.
- Smith, S. J., D. S. Burch, and L. M. Branscomb, *Ann. Geophys.*, **14**, 225-231, 1958.

(Manuscript received January 12, 1961.)

The Relationship of Low-Height Ionosonde Echoes to Auroral-Zone Absorption and VHF D Scatter

JENS K. OLESEN

*Danish National Committee, URSI
Danmarks Tekniske Højskole, Copenhagen, Denmark*

AND

J. W. WRIGHT

*Central Radio Propagation Laboratory, National Bureau of Standards
Boulder, Colorado*

Abstract. Weak diffuse HF reflections at heights between 75 and 95 km in the auroral zone are designated $E_s d$. Diurnally, $E_s d$ occurs most frequently at noon at all seasons, and shows a pronounced minimum of occurrence around 2000 hours local time. There is little evidence for a seasonal variation from the year's data presented here. It is suggested that $E_s d$ is related to the anomalies of auroral-zone absorption and that it is the layer responsible for VHF forward scatter.

Introduction. Ionospheric soundings using a J-3 type ionosonde (peak power 10 kw, 1–25 Mc/s, vertical half-rhombic antennas), at Narsarsuaq, Greenland (61.2°N, 45.4°W; 71.2°N geomag.), operated in cooperation with the National Bureau of Standards by the Danish National Committee of URSI, very often show diffuse gain-sensitive echoes at heights between 75 and 95 km in the frequency range 2 to 8 Mc/s. Owing to the frequent and characteristic appearance of this echo, it has been given specific attention in the systematic reduction of data from this station since September 1958. In accordance with provisions by URSI¹ for the description of new *E*-region phenomena, we have designated this echo $E_s d$.²

An ionogram illustrating $E_s d$ is shown in Figure 1. Characteristically, $E_s d$ appears as a weak, rather diffuse reflection at low ionospheric

heights, usually in the range 75 to 95 km. The height of reflection is independent of frequency, but the echoes often appear to be returned from a range of altitudes, in contrast to the occasional meteor echoes and other, more rare, reflections from this height range. The layer is remarkably reliable in its occurrence. Similar echoes are observed at other auroral-zone stations (for example, College, Alaska, and Godhavn, Greenland), but the preliminary results reported here are derived solely from observations made at Narsarsuaq, Greenland.

The diurnal, seasonal, and height variability of $E_s d$. The diurnal and seasonal occurrences of $E_s d$ at Narsarsuaq, from a year's data (September 1958–August 1959) are illustrated in Figure 2. At all seasons, there is one daily principal maximum of occurrence at noon, and sometimes a lower peak in the morning. A deep minimum is evident in the early evening (17–23 hours local time) at all seasons. There is a suggestion of a summer maximum, but further data are needed to confirm it. On the whole, it appears from our data that these echoes are visible on at least 70 per cent of the noon ionograms, at all seasons.

Selecting January 1959 as a typical month for more intensive study, the diurnal height behavior of $E_s d$ is illustrated in Figure 3. The height is minimum at noon, the median value being 85

¹ See the discussion of E_s types in the second report of the URSI special committee on worldwide ionospheric soundings (URSI information bulletin 99, pp. 48–90) or the discussion in *Ann. IGY*, vol. III, part 1, p. 94, for a description of the E_s typing system.

² It could be maintained that this phenomenon might better be termed 'sporadic D'; however, it appears that modern usage of the designation ' E_s ' refers to all unusual or nonregular occurrences in the lower ionosphere, as distinct from the comparatively well-understood *E* layer.

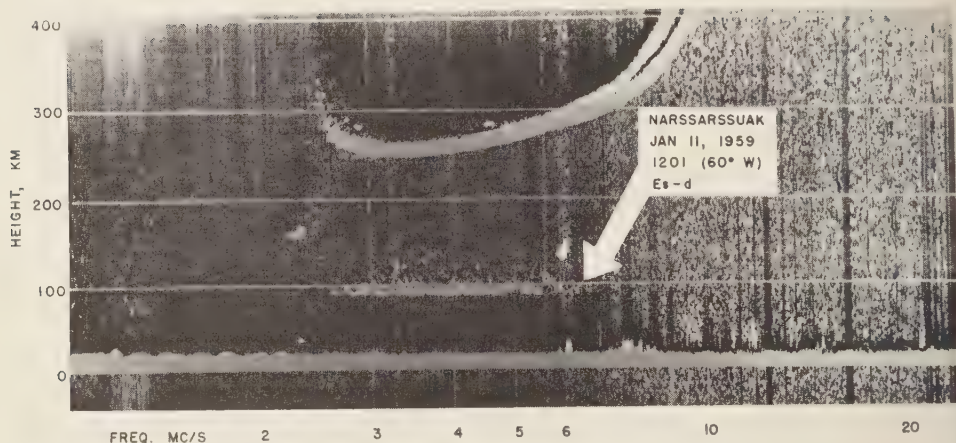


Fig. 1. Typical example of $E_s d$; an ordinary-day observation.

km. In the late morning and early evening hours the median height increases to 90 and 93 km, respectively. Thus, in the hours when $E_s d$ is most frequent, it is also observed at lower heights. In the late evening and near midnight, $E_s d$ occurs too infrequently to permit reliable conclusions about its height at these hours.

The height of $E_s d$ varies somewhat irregularly throughout the day, changes of 5 to 10 km being common. This is illustrated in Figure 4, the height vs. time being plotted (continuous line) for several days of January 1959 in which $E_s d$ was reliably observed.

$E_s d$ and auroral-zone absorption. It is interesting to note that the diurnal occurrence frequency of $E_s d$ is very similar to the diurnal variation of absorption in the auroral zone. Little and Leinbach [1958] have reported 1 month (March) of riometer (cosmic-noise absorption) data from College and Barrow, Alaska. These results show the same noon peak and evening minimum as observed in the occurrence of $E_s d$. Little and Leinbach place the height of the absorbing region as between 60 and 100 km.

We have attempted to examine the relationship between the occurrence of $E_s d$ and absorption by comparing the measurements of f_{\min} (minimum radio frequency of E - or F -region echo on the ionogram, due to absorption) with the height of $E_s d$. For the January data, the measured value of f_{\min} is plotted vs. $h'E_s d$ in Figure 5. It is clear that the higher values of f_{\min} (higher

absorption) are correlated with lower $E_s d$ heights.

It is noteworthy that at times of intense absorption (blackout conditions, implying no radio echoes observed from the E and F regions), $E_s d$ reflections may frequently be observed at unusually low heights, often down to 70 km. This suggests that at such times the $E_s d$ reflections return from the lower part of the height region where most of the absorption is taking place. An example is shown in Figure 6 from the ionospheric storm of March 26, 1958, where $E_s d$ is observed at a height of 74 km during blackout.

The relation between absorption and $h'E_s d$ is apparent in the individual observations also. In Figure 4, f_{\min}^F or f_{\min}^E (dashed line) is seen to vary closely with $h'E_s d$ on days selected for continuity and clarity of the $E_s d$ layer.

$E_s d$ and the ionospheric 'VHF scattering region.' Five of the distinctive characteristics of $E_s d$ —its height at about 85 km, the pronounced diurnal minimum of occurrence at 17 to 22 hours local time, the diurnal maxima in the early morning and at noon, the summer maximum, and its enhancement during disturbed conditions—lead us to suggest that this layer is identical with the region responsible for ionospheric D scattering at VHF studied by Bailey, Bateman, and Kirby [1955]. At all seasons, and in middle and high latitudes, the diurnal variations in VHF signal strength follow closely the occurrence of $E_s d$ in the auroral zone. It is significant to note that the VHF D -scatter signal is always found to be

present, at all latitudes and seasons; it has also been shown to be stronger and with a larger diurnal variation at high latitudes than at moderate latitudes. Similarly, the $E_s d$ reflections are observed at noon at least 70 per cent of the time at Narssarssuak, and are present at all seasons. Additional evidence on the diurnal varia-

tion of VHF D scatter has been given by Hagfors [1957].

The height of the VHF D scattering layer has been inferred from observations of signal strength variations vs. path length for paths near the maximum possible for one-hop propagation. Bailey and co-workers conclude that the day-

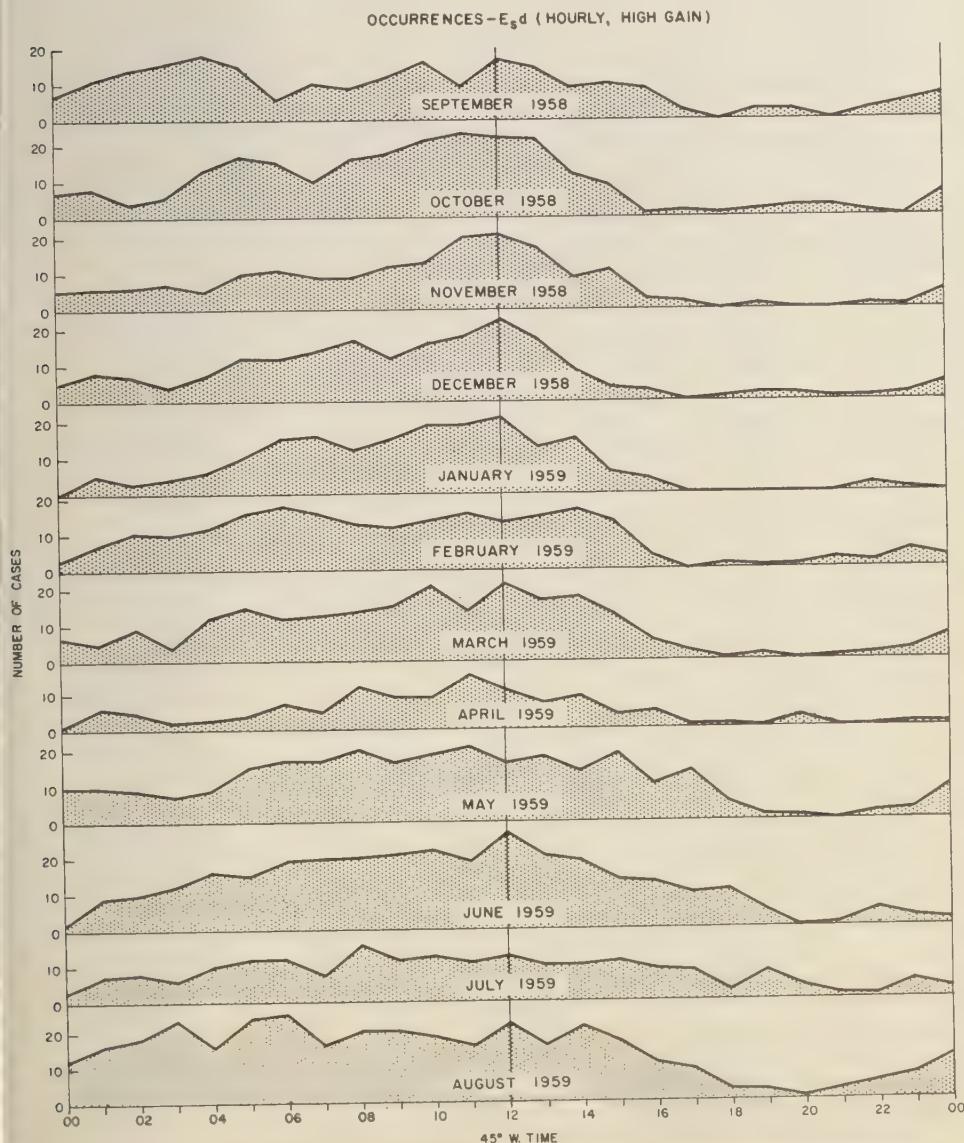


Fig. 2. Diurnal and seasonal occurrence of $E_s d$ at Narssarssuak, September 1958 to August 1959.

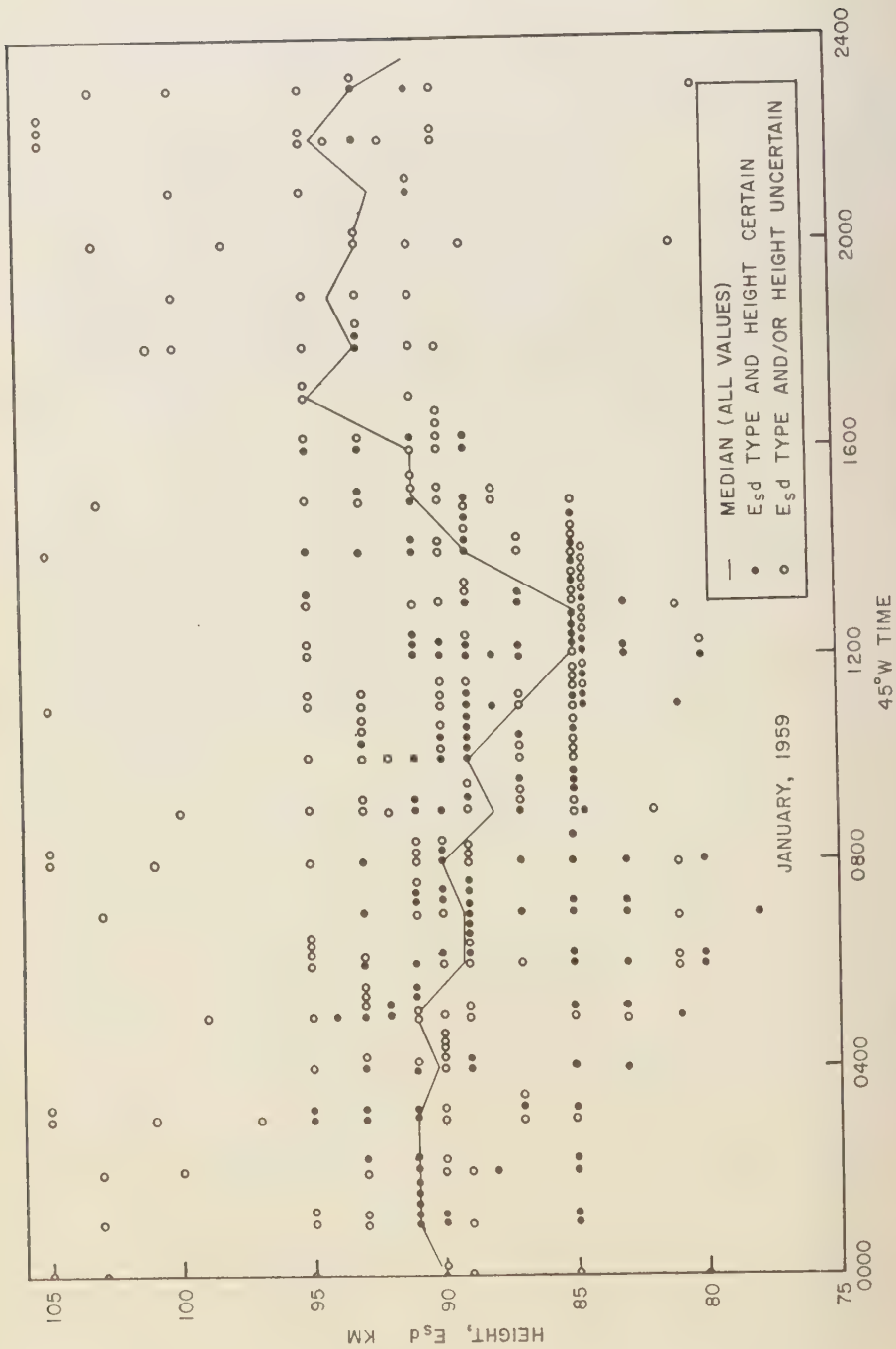


Fig. 3. Diurnal variation of $E_s d$ height, January 1959 data at Narsarsuaq.

time heights lie between 75 and 85 km, and the nighttime heights in the range from 85 to 90 km. This would appear to be in good agreement with the lower limits of the diurnal height variation illustrated in Figure 3. Finally, the initial enhancement of VHF scatter signals during polar blackout conditions is consistent with our observation that $E_s d$ is usually lower and more strongly reflecting in these conditions.

Bailey and his co-workers attempted to measure directly the height of this layer by vertical-incidence 1000-kw pulse measurements at 50 Mc/s; they failed to observe echoes. The $E_s d$ observations suggest the reason for this negative result: Evidence has been given [Wright and Lautier, 1960] to suggest that E_s echoes approximately obey the 'equivalence theorem,' relating the maximum radio frequency returned at verti-

cal incidence by a stratum to the maximum returned at a larger angle. By this theorem, over a 1500-km path, a stratum at 80 km will propagate a radio wave of frequency roughly 6 times larger than at vertical incidence. Thus, since most of the oblique-incidence VHF observational evidence has been collected at frequencies near 50 Mc/s, the equivalence theorem would suggest that vertical-incidence measurements should be attempted on frequencies no greater than about one-sixth of this. Our values of $fE_s d$, which rarely exceed 8 Mc/s, are consistent with this conclusion. More recently, Bowles [1958] has reported pulse observations at 41 Mc/s, using a system of very high gain (peak pulse power 4-6 megawatts, 35 db/isotropic antenna gain) in which it appears that vertical-incidence returns are obtained.

$E_s d$ and other low-height echoes observed by pulse

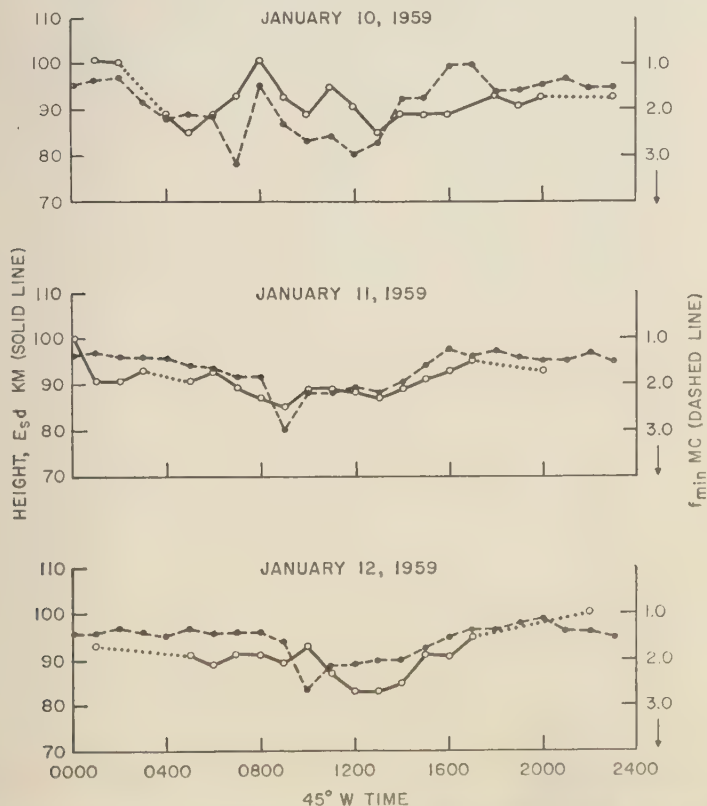


Fig. 4. Diurnal variation of $E_s d$ height and f_{\min} on three ordinary days, suggesting that $E_s d$ height is closely related to absorption.

methods. Vertical-incidence pulse observations at 1.75 Mc/s have been made at moderate latitudes in New Zealand by Gregory [1956]; and the same author [Gregory, 1957] has suggested the similarity of his 80- to 90-km layer to the VHF scattering level. The New Zealand data were obtained at a geomagnetic latitude of 48°S, where a definite seasonal variation in height was observed. (Dieminger [1958] reported similar observations on 1.65 Mc/s at Lindau; a winter maximum of occurrence is evident in his data.) Although it may be noted that no significant seasonal variation in occurrence of $E_s d$ is detected in our Narssarssuak data the height data are as

yet insufficient to examine a possible seasonal height variation.

Reports of low-height echoes at auroral latitudes have been made by Landmark [1958], Stoffregen [1958], and Lindquist [1951]. The first work describes the occurrence of these echoes during polar blackouts, but notes that such echoes appear to be present continuously when specifically looked for. Landmark interprets the echoes during polar blackouts as coming from the lower boundary of the region causing the enhanced 'polar-blackout' absorption.

Stoffregen [1958] has reported echoes at low frequencies (0.33–1.5 Mc/s) in the height range

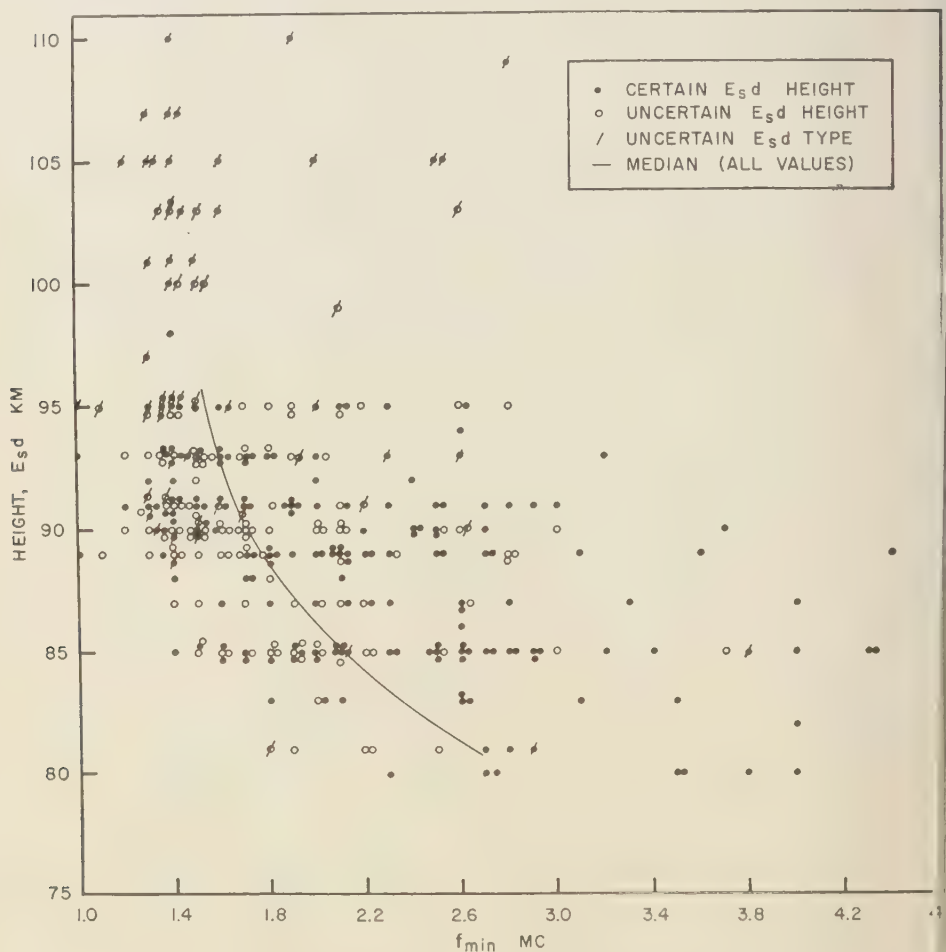
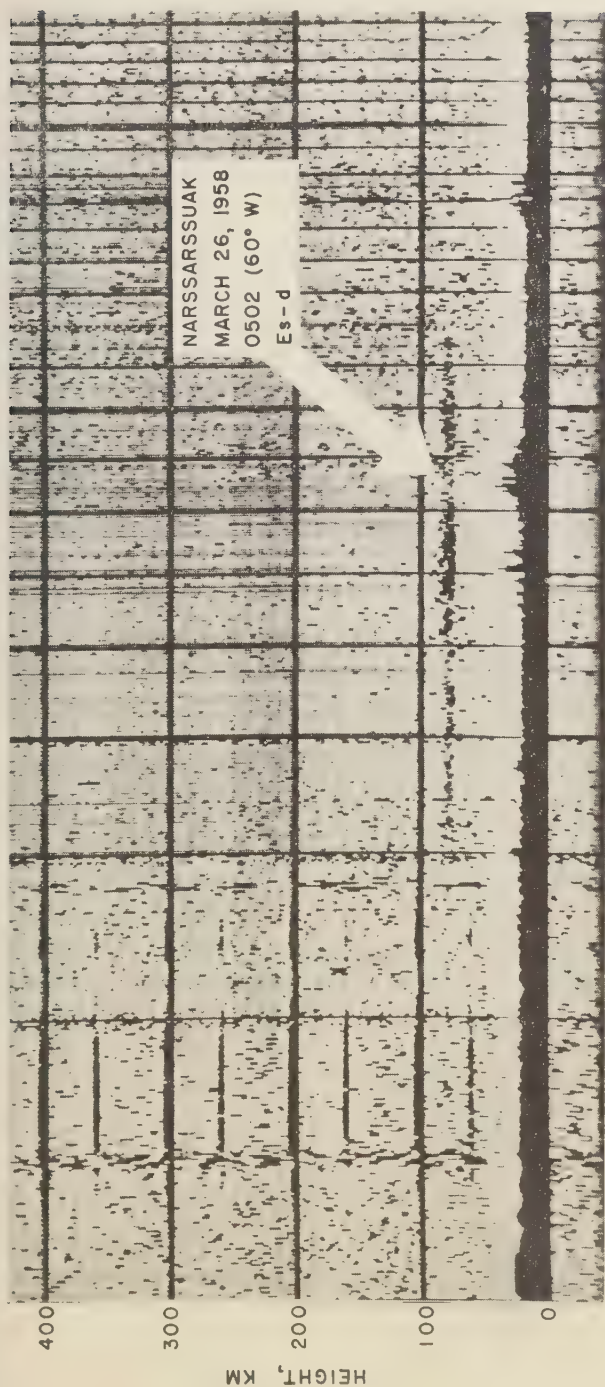


Fig. 5. Height of $E_s d$ vs. f_{min} , January 1959 data.

FREQ. MC/S 2 3 4 5 6 10 20



FREQ. MC/S 2 3 4 5 6 10 20

Fig. 6. $E_s d$ observed during polar blackout conditions. Note low height of reflection at 74 km.

75 to 90 km during intense auroral activity; Lindquist's [1951] results are similar.

Summary. Summarizing our own preliminary conclusions together with those of the authors cited above, it is evident that a region capable of returning vertically incident high-frequency signals exists at heights within the range 75 to 95 km with the following properties: (1) It is present at moderate and auroral latitudes, at all seasons and all times of day. (2) The occurrence of auroral-zone absorption is closely connected with the existence of this layer, there being evidence that during periods of high absorption these reflections come from a height *below* the region of strongest absorption. (3) The low-height echoes are enhanced considerably during polar blackouts and auroral activity. (4) The reflecting region is probably identical with the level responsible for VHF *D* scatter; it appears that measurements such as ours provide direct vertical-incidence observations of reflection height in the auroral zone.

REFERENCE

- Bailey, D. K., R. Bateman, and R. C. Kirby, *Proc. IRE*, **43**, 1181, 1955.
 Bowles, K. L., *Phys. Rev. Letters*, **1**, H523, 1958.
 Dieminger, W., *Report of the Physical Society Conference on the Physics of the Ionosphere, Cavendish Laboratory, Cambridge, September 1954*, p. 53, Physical Society, 1958.
 Gardner, F. F., and I. L. Pawsey, *J. Atmospheric and Terrest. Phys.*, **3**, 321, 1953.
 Gregory, J. B., *Australian J. Phys.*, **9**, 324, 1956.
 Gregory, J. B., *J. Geophys. Research*, **62**, 383, 1957.
 Hagfors, T., *Polar Atmosphere Symposium*, Pergamon Press, pt. 2, p. 205, 1957.
 Landmark, B., *J. Atmospheric and Terrest. Phys.*, **12**, 99, 1958.
 Lindquist, R., *Chalmers Tek. Högskol. Handl.*, no. 103, 1951.
 Little, G. C., and H. Leinbach, *Proc. IRE*, **46**, 334, 1958.
 Stoffregen, W., *J. Atmospheric and Terrest. Phys.*, **13**, 167, 1958.
 Wright, J. W., and T. N. Gautier, *J. Research NBS-D*, pp. 347-8, 1960.

(Manuscript received January 7, 1961.)

Photogrammetric Refraction Angle: Satellite Viewed from Earth

B. L. JONES

Midwest Research Institute
Kansas City, Missouri

Abstract. If an earth satellite is photographed from the earth against a star background for geodetic purposes, an accurate determination of the photogrammetric refraction angle ρ must be made. It is shown that ρ is very sensitive to the astronomical refraction angle R_∞ for satellite heights up to at least 10^3 km if the satellite zenith-distance is less than about 70° . For example, the error in ρ contributed by an error in R_∞ is 31 times the error in R_∞ for a satellite at a height of 200 km viewed at a zenith-distance of 10° .

Introduction. Recent considerations of the use of artificial earth satellites and rockets as triangulation points for geodetic purposes suggested a study of refraction of light along the path between an extraterrestrial source and an earth-based observation station. A relationship between the photogrammetric refraction angle ρ and its determining variables is derived, and the sensitivities of ρ to these variables is determined and analyzed.

Photogrammetric refraction angle. In the derivation of the refraction equation, two assumptions are made: (1) the atmosphere has a spherically stratified index of refraction along the light path and (2) the refractive index at the satellite and between the satellite and the stars is unity.

Snell's law for a light path in a spherically stratified medium is

$$mr \sin i = k \quad (1)$$

where

m is the refractive index at any point on the path.

r is the distance between the above point and the center of curvature of the stratification.

i is the angle between the tangent to the path at the above point and the radius vector at the point (the incidence angle).

k is a constant, independent of the location of the point on the path.

It follows from (1) and the first part of (2) that

$$i_1 = \sin^{-1} [(m_0 r_0 \sin i_0) / r_1] \quad (2)$$

where the subscript 0 refers to the observer

location and the subscript 1 refers to the satellite location. The geometry is shown in Figure 1.

To an observer at P , a star on the straight line extension of the light path, indicated with an arrow in Figure 1, has the same apparent direction PT as the satellite. The true direction QS of this fictitious star may be obtained by standard astrometric procedures from the photograph of the satellite against the background of stars whose celestial coordinates are known. This direction differs from the desired direction PS by the angle ρ , referred to here as the photogrammetric refraction angle.

From the triangle QOS in the figure,

$$\theta = i_0 + R_\infty - i_1 \quad (3)$$

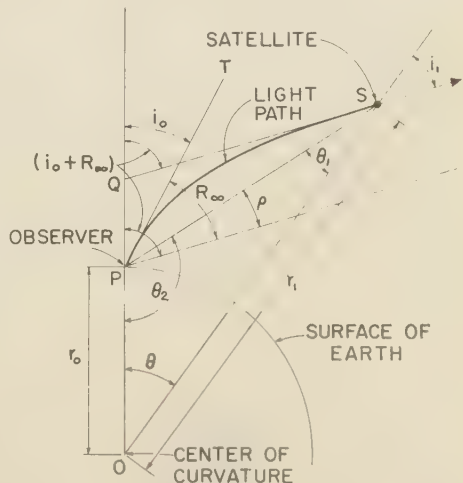


Fig. 1. Refraction geometry for satellite outside of optically refracting atmosphere.

where R_∞ is the astronomical refraction angle. Applying the law of tangents to triangle POS yields

$$\frac{1}{2}(\theta_2 - \theta_1) = \tan^{-1} \left[\frac{r_1 - r_0}{r_1 + r_0} \tan \frac{1}{2}(\theta_2 + \theta_1) \right] \quad (4)$$

From the figure,

$$\frac{1}{2}(\theta_2 + \theta_1) = \frac{1}{2}(\pi - \theta) \quad (5)$$

and

$$\rho = R_\infty - \pi + i_0 + \theta_2 \quad (6)$$

Combining (4), (5), and (6) and simplifying yields

$$\rho = R_\infty - \frac{\theta}{2} + i_0 - \tan^{-1} \left(b \tan \frac{\theta}{2} \right) \quad (7)$$

where

$$b = (r_1 + r_0)/(r_1 - r_0)$$

Inspection of (2), (3), and (7) indicates that ρ may be considered a function of four variables,

$$\rho = \rho(R_\infty, m_0, r_0/r_1, i_0) \quad (8)$$

For the purpose of this discussion, however, ρ will be considered a function of five variables,

$$\rho = \rho(R_\infty, m_0, r_0, r_1, i_0) \quad (9)$$

The change in ρ resulting from small changes in the variables is

$$\begin{aligned} \Delta\rho = & \frac{\partial\rho}{\partial R_\infty} \Delta R_\infty + \frac{\partial\rho}{\partial m_0} \Delta m_0 \\ & + \frac{\partial\rho}{\partial r_0} \Delta r_0 + \frac{\partial\rho}{\partial r_1} \Delta r_1 + \frac{\partial\rho}{\partial i_0} \Delta i_0 \end{aligned} \quad (10)$$

The partial derivatives obtained from (7) are

$$\begin{aligned} \partial\rho/\partial R_\infty = & \{1/2\} - \{[b/2]/[\cos^2(\theta/2) \\ & + b^2 \sin^2(\theta/2)]\} \equiv \alpha \end{aligned} \quad (11)$$

$$\partial\rho/\partial m_0 = (1 - \alpha)(\tan i_1)/m_0 \equiv \beta \quad (12)$$

$$\begin{aligned} \partial\rho/\partial r_0 = & \{m_0\beta/r_0\} \\ & + \{[r_1(2\alpha - 1) \sin \theta]/[r_1^2 - r_0^2]\} \equiv \gamma \end{aligned} \quad (13)$$

$$\partial\rho/\partial r_1 = -\gamma r_0/r_1 \quad (14)$$

$$\partial\rho/\partial i_0 = \alpha + (\beta m_0/\tan i_0) \quad (15)$$

Table 1 shows values of ρ and its partial derivatives for various values of the variables. The satellite height is $(r_1 - r_0)$. The partial derivative $\partial\rho/\partial m_0$ is expressed in seconds per n -unit where one n -unit is one unit in the sixth decimal of m_0 .

Perhaps the most interesting fact revealed by Table 1 is that, if there are no errors in m_0 , r_0 , r_1 , and i_0 , the error $\Delta\rho$ is, in general, many times the error ΔR_∞ . For example, at $i_0 = 30^\circ$ and $r_1 - r_0 = 400$ km, $\Delta\rho = -12.06 \Delta R_\infty$. This may or may not be a serious error in ρ . If it is, it is likely that the calculated value of R_∞ is in error by less than, say, $0.002''$ at $i_0 = 30^\circ$, then it is likely that the error in ρ is less than $0.024''$ for $r_1 - r_0 = 400$ km. This error may be tolerable in geodetic work, but an error several times as large may not be.

For a spherically stratified atmosphere, the total astronomical refraction is [Smart, 1956]

$$R_\infty = r_0 m_0 \sin i_0 \int_1^{m_0} \frac{dm}{m(r^2 m^2 - r_0^2 m_0^2 \sin^2 i_0)^{1/2}} \quad (16)$$

the limits of integration being m_0 at the observer and unity at the highest layer. Although the value of R_∞ is primarily dependent on the values of r_0 , m_0 , and i_0 , it is also dependent on the relation between m and r . This relation will vary with time and with observer location. It has been found by astronomers that values of R_∞ may be approximated by using (16) and a model atmosphere based only on the conditions at the observation point and on assumed conditions for $r > r_0$. The refraction equation (21) of Willis [1941], for example, is based on these assumptions. This equation is among the best available, but it cannot yield exact values of R_∞ even if r_0 , m_0 , and i_0 are exact.

A literature search has not uncovered any experiment or analysis that would allow a guess at the accuracy of any modern astronomical refraction equation except for i_0 values greater than about 80° . Strand [1953] concluded from his observations at Dearborn Observatory that on individual nights the actual astronomical refraction may differ as much as 8 per cent from values computed from tables in *Astronomisch-Geodatisches Jahrbuch* (1952). Although no

¹ $R_\infty = \tan i_0 (\log_e m_0) f$, where the values of the function f are obtained from Willis's tables.

specifically stated, this difference apparently applies for $80^\circ < i_0 < 90^\circ$. Lacking information of the probable errors of astronomical refraction tables for $i_0 < 80^\circ$, it is impossible to estimate for $i_0 < 80^\circ$ the probable error in ρ .

In Willis's equation 21 [1941] R_∞ is shown (inexactly) as a function only of m_0 , r_0 , i_0 , and the temperature, t_0 , at the observation point. If i_0 is less than about 70° and r_1 is fixed, nearly the same computed value of ρ will result from

TABLE 1. Computed values of ρ and its Partial Derivatives for $m_0 = 1.0002818$ and $r_0 = 6,370.06$ km

i_0 , deg	R_∞ , sec	$r_1 - r_0$, km	ρ , sec	$-\partial\rho/\partial R_\infty$, sec/sec	$\partial\rho/\partial m_0$, sec/n-unit	$\partial\rho/\partial r_0$, sec/km	$-\partial\rho/\partial r_1$, sec/km	$\partial\rho/\partial i_0$, sec/min
0	0.0	100	0	63.70	0	0	0	1.077
		200	0	31.85	0	0	0	0.538
		400	0	15.93	0	0	0	0.269
		10^3	0	6.37	0	0	0	0.108
		10^4	0	0.64	0	0	0	0.011
		10^5	0	0.06	0	0	0	0.001
		∞	0	0	0	0	0	0
10	10.2	100	2.89	61.79	2.248	0.0276	0.0272	1.078
		200	1.53	30.90	1.124	0.0078	0.0076	0.539
		400	0.76	15.46	0.562	0.0020	0.0019	0.269
		10^3	0.30	6.19	0.225	0.0003	0.0003	0.108
		10^4	0.04	0.62	0.023	0.0000	0.0000	0.011
		10^5	0.00	0.06	0.002	0.0000	0.0000	0.001
		∞	0	0	0	0	0	0
30	33.4	100	7.34	47.89	5.705	0.0866	0.0853	1.077
		200	3.86	24.00	2.859	0.0198	0.0192	0.540
		400	1.94	12.06	1.436	0.0051	0.0048	0.271
		10^3	0.78	4.88	0.582	0.0011	0.0009	0.110
		10^4	0.08	0.52	0.062	0.0000	0.0000	0.012
		10^5	0.01	0.05	0.007	0.0000	0.0000	0.001
		∞	0	0	0	0	0	0
50	68.9	100	10.08	26.59	6.540	0.1019	0.1003	1.082
		200	5.19	13.43	3.302	0.0262	0.0254	0.546
		400	2.65	6.84	1.682	0.0068	0.0064	0.278
		10^3	1.03	2.86	0.705	0.0011	0.0010	0.117
		10^4	0.14	0.35	0.087	0.0000	0.0000	0.014
		10^5	0.02	0.04	0.010	0.0000	0.0000	0.002
		∞	0	0	0	0	0	0
70	157.8	100	16.10	7.836	4.451	0.1574	0.1549	1.111
		200	8.50	4.099	2.328	0.0401	0.0389	0.582
		400	4.57	2.206	1.253	0.0106	0.0099	0.314
		10^3	2.12	1.024	0.582	0.0020	0.0017	0.146
		10^4	0.34	0.166	0.095	0.0001	0.0000	0.024
		10^5	0.04	0.021	0.012	0.0000	0.0000	0.003
		∞	0	0	0	0	0	0
80	317.3	100	31.70	2.286	2.697	0.275	0.271	1.209
		200	17.53	1.290	1.522	0.073	0.071	0.686
		400	10.52	0.759	0.896	0.021	0.019	0.406
		10^3	5.49	0.396	0.467	0.004	0.004	0.213
		10^4	1.09	0.078	0.092	0.000	0.000	0.042
		10^5	0.15	0.010	0.012	0.000	0.000	0.006
		∞	0	0	0	0	0	0
90	2,059.5	100	367.36	-0.0440	1.119	1.812	1.784	2.642
		200	261.71	-0.0333	0.793	0.665	0.645	2.000
		400	185.34	-0.0246	0.560	0.248	0.233	1.477
		10^3	115.66	-0.0159	0.349	0.070	0.061	0.956
		10^4	28.77	-0.0041	0.087	0.005	0.002	0.248
		10^5	4.10	-0.0006	0.012	0.001	0.000	0.036
		∞	0	0	0	0	0	0

(7) for a group of values of R_∞ , m_0 , r_0 , i_0 , and t_0 that satisfy Willis's equation precisely as will result if the values of any number of these five variables are changed considerably and in arbitrary directions, provided that the new values also satisfy Willis's equation precisely. Thus, if Willis's equation is used, there is no need to determine m_0 , r_0 , i_0 , and t_0 accurately in order to compute ρ . If i_0 is less than about 70° , the respective errors can be as much as several n -units, a few kilometers, a few minutes, and several degrees Centigrade without changing the resulting computed value of ρ by more than a few hundredths of a second of arc from the value that would be computed if there had been no errors in m_0 , r_0 , i_0 , and t_0 .² But this change in the computed value of ρ is not the error in ρ ; the error is the difference between the actual and the computed value. As has already been indicated, the error cannot be estimated without knowledge of the accuracy of the astronomical refraction equation used.

Conclusions. It is concluded from this study that (1) the error in the computed photo-

² The groups of values of R_∞ , m_0 , r_0 , and i_0 selected for Table 1 do not satisfy Willis's equation precisely for reasonable values of t_0 . A decrease in m_0 of about 0.7 n -units (more or less, depending on i_0) would make the selected quantities satisfy the equation for reasonable t_0 values.

grammetric refraction angle can vary from more than 60 times the error in the astronomical refraction angle determination for satellites at low heights to less than the error in the astronomical refraction angle determination for satellites at heights greater than a few thousand kilometers and (2) the probable errors in modern astronomical refraction equations do not appear to be known.

Acknowledgments. I am indebted to Leon Findley, who suggested the investigation, and to Wanda Chinnery and Betty Kahn, who calculated the data in Table 1.

This work was performed under contract DA 44-009-ENG-3769, project 8T35-12-001-01, for U. S. Army Engineer Geodesy, Intelligence and Mapping Research and Development Agency, Fort Belvoir, Virginia.

REFERENCES

- Smart, W. M., *Text-Book on Spherical Astronomy*, Fourth Edition, The University Press Cambridge, England, 430 pp., p. 64, 1956.
- Strand, K. G., Investigation of atmospheric refraction at low altitudes, *ADTIC Publ. No. A-102, AD-6452*, Maxwell Air Force Base, Alabama, 1953.
- Willis, J. E., A determination of astronomical refraction from physical data, *Trans. Am. Geophysical Union*, 22, 324-336, 1941.

(Manuscript received December 10, 1960.)

The Diurnal Variation of K Indices of Geomagnetic Activity on Quiet Days in 1940–1948

SETH B. NICHOLSON

*Mount Wilson and Palomar Observatories, Carnegie Institution of Washington
California Institute of Technology, Pasadena, California*

AND

OLIVER R. WULF

*United States Weather Bureau, California Institute of Technology
Pasadena, California*

Abstract. Local- and universal-time components of the diurnal variation of the K indices have been computed from the eight daily K numbers for the 10 quiet days of each month in the 9 years 1940–1948 from six observatories in moderately low latitudes and fairly evenly distributed in longitude. Because the quiet days were universal time days, the universal time component contains a statistical curvature effect. The magnitude of this has been estimated. The universal time component is less regular than the local time component. When the local time component is examined for each of the 9 years separately, a pronounced change in form of much apparent regularity is exhibited over this 9-year interval. While it is possible that the variability of the amplitude and form of the quiet-day daily variation of the field contributed appreciably to the K numbers for these quiet days, it seems probable that this change of form of the local-time component, particularly marked from 1943 to 1947, arises mainly from disturbance and represents a change in the relative amounts of disturbance having different forms of diurnal variation. It may represent a change of this distribution with the solar cycle.

It seems clear that winds in the ionosphere may contribute in some measure to geomagnetic disturbance by generating, through dynamo action, electric currents in addition to those that lead to the quiet-day daily variation of the earth's field. One property of geomagnetic disturbance that may contain evidence relevant to the part played by winds in its production is its daily variation.

It has seemed to us of value, therefore, to study the diurnal variation of the irregular geomagnetic fluctuations, and the 3-hour-range index K affords one means of doing this. In studying the diurnal variation of the K numbers for six rather low latitude observatories [Nicholson and Wulf, 1955; 1958], a local and a universal time component were obtained using all the days.

In an effort to bring out the diurnal features characteristic of disturbance, separate results have now been obtained for the 10 quiet, for the 5 disturbed, and for the remaining intermediate days of each month, as well as for all days. The present paper gives the results for the 10 quiet days of each month.

The eight daily K numbers for the 10 quiet days were used for the 9 years 1940–1948 [Johnston, Scott, and Balsam, 1948; Howe and Weisman, 1949] from the six observatories—Alibag, Watheroo, Honolulu, Tucson, San Juan, and San Fernando—which are fairly evenly distributed in longitude. The 9-year-average departures of the K numbers for each of the eight 3-hour intervals were computed for each observatory for each month of the year. The 12-month-average departures for each of the eight intervals were also computed for each observatory for each of the 9 years.

These departures were arranged on local time and averaged, giving a local-time component of the diurnal variation. This local-time component was subtracted from each of the six sets of original departures. These differences were then arranged on universal time and averaged, giving a universal-time component that was used to obtain a second approximation of the local- and the universal-time components, in order to minimize the effects of the unevenness in the distribution in longitude of the stations used.

Station residuals, as described in *Nicholson and Wulf* [1958], were, of course, present.

The 9-year average local-time component of the diurnal variation of K for the 10 quiet days of each month is shown in Figure 1. The unit used throughout is 0.01 K . These results indicate an appreciable diurnal variation on local time of considerable regularity. No pronounced seasonal change is apparent, though there are differences such as the accentuation of the 9–15-hour maximum in June and July.

At the foot of Figure 1, the average of the 12 monthly variations for these 9 years is shown. The standard deviations of these eight 12-month-average values were small, the largest being 1.6 units (0.016 K) in 21h and in 0h.

The radius of the dots used to indicate the points is approximately 2.5 units. These standard deviations contain, of course, any real seasonal changes as well as the random fluctuations of the data. Thus, the twelve 9-year-average monthly variations resembled their average fairly closely. Figure 3, however, shows that the nine individual yearly averages, of which the curve at the foot of Figure 1 is also the average, changed conspicuously over these 9 years.

We have also carried through the same treatment for 6 months of each of the 9 years for the 5 quiet days, which are the quieter 5 of the 10. The months chosen for this study were January, March, June, July, September, and December. The monthly results resembled those of Figure 1,

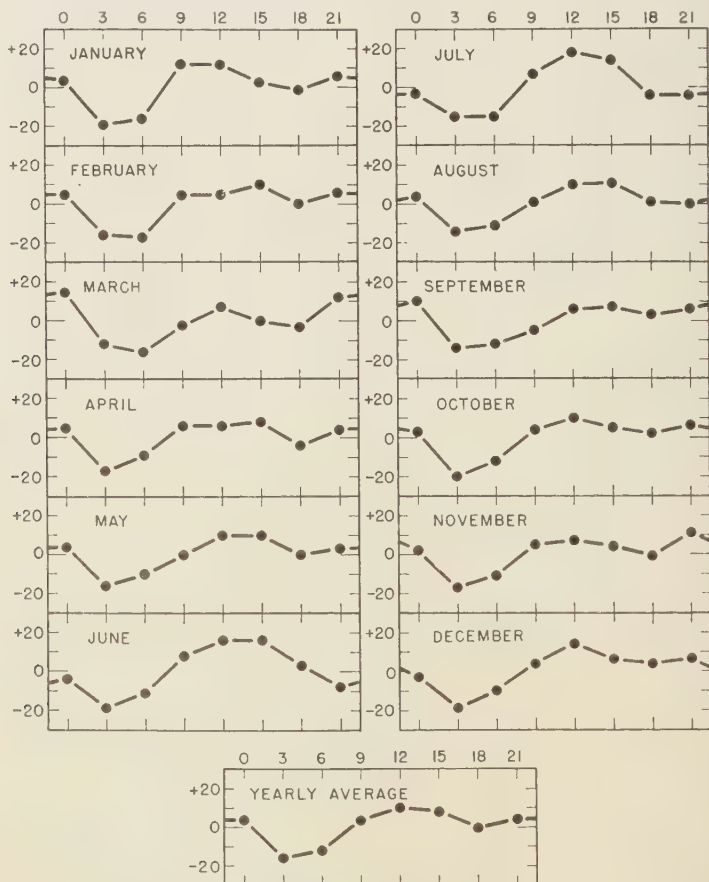


Fig. 1. Local-time component of the diurnal variation of K , 1940–1948, for the 10 quiet days of each month from the six observatories. Unit, 0.01 K .

indicating that if the principal features of Figure 1 were introduced by disturbance, the 5 quiet days contained an amount comparable at least with the 10 quiet days.

Figure 2 shows the 9-year-average universal-time component of the diurnal variation for each month, and curve C in the lower right-hand corner shows the average of the 12 monthly variations. The monthly behavior is less regular throughout the year than that of the local-time component and the amplitude of the yearly average is smaller.

The universal-time component for the selected quiet days contains a statistical curvature effect

[*Bartels and Johnston*, 1939, see p. 467, sec. 7], since these days were selected on the basis of the Greenwich (UT) day. An estimate of this curvature was obtained for the months of June and September as samples. This was done by tabulating for each station not only the eight K numbers for the 24 hours of the UT quiet day but also the four values for the 12 hours preceding that day and the four values for the 12 hours following that day. This work yielded the average K numbers for each station for sixteen consecutive 3-hour intervals centered on the 12th hour of the average quiet day. Eight interval (24-hour) running means were then obtained centered on

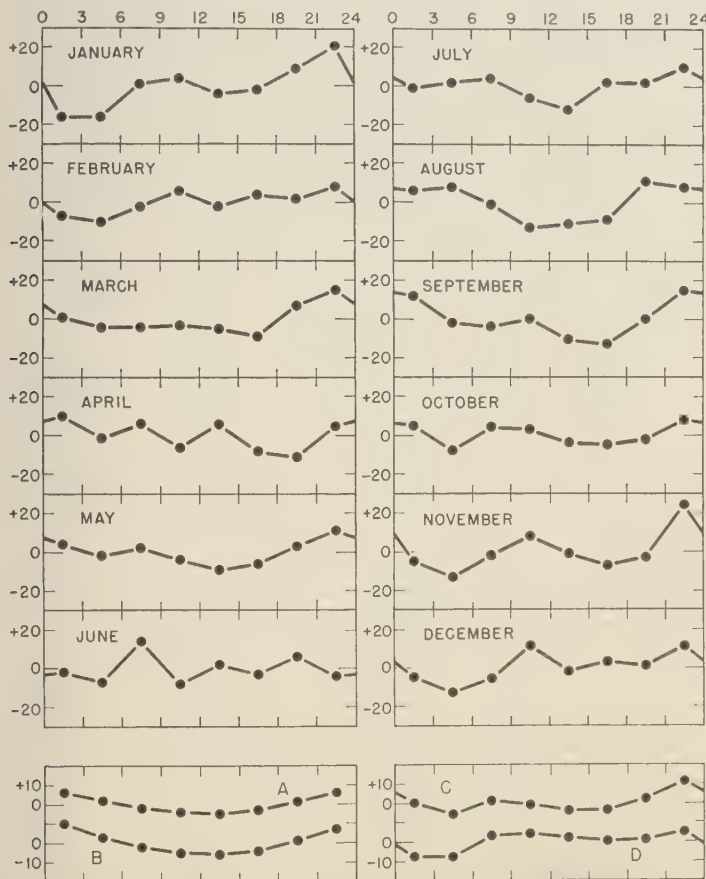


Fig. 2. Universal-time component of the diurnal variation of K , 1940-1948, for the 10 quiet days of each month after subtracting the local-time component. C is the yearly average. A and B illustrate statistical curvature for June and September, respectively. D is C minus the average of A and B . Unit, 0.01 K .

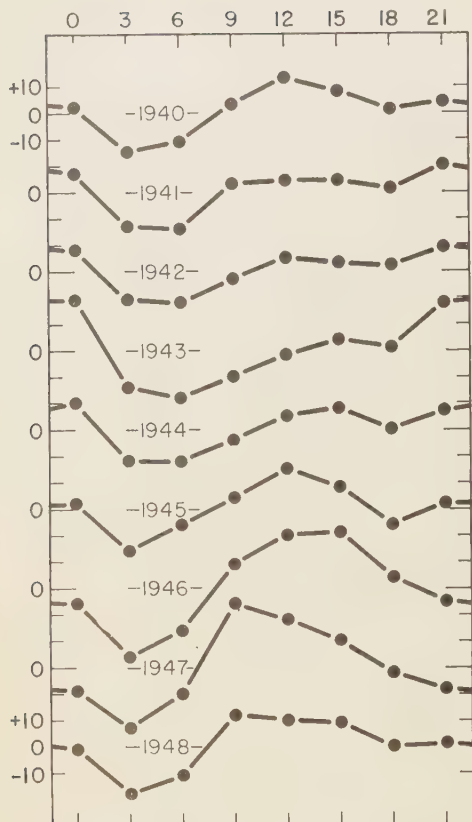


Fig. 3. Yearly average local-time component of the diurnal variation of K , 1940-1948, for the 10 quiet days of each month from the six observatories for each of the 9 years. Unit, 0.01 K .

each of the 9 hours (0, 3 . . . 21, 24) of the average quiet day. In each of these means the daily variation should be averaged out. The average of the means at the beginning and end of each 3-hour interval was taken as the value for the interval.

The eight values so obtained for each station were averaged, and their departures taken, giving for the months of June and September estimates of the statistical curvature in this 10-quiet-day data. The six (station) sets were similar for both months, each exhibiting the expected concave-upward curvature. The six estimates of curvature, averaged for each of these 2 months, gave the results shown in the lower left-hand portion of Figure 2; curve A is

for June and B for September. The standard deviations of these average values were computed from the six values of which each is the average. The largest standard deviation was 1.5 units in the 0-3-hour interval in the data for June. The dots used in the figure have a radius of approximately 2.5 units. Curve D in the lower right-hand corner of Figure 2 is curve C minus the average of curves A and B, that is, the yearly average universal-time component corrected approximately for curvature.

Although its amplitude is smaller, this universal-time component corrected for curvature resembles that for all days, which, of course, needs no curvature correction [Nicholson and Wulf, 1958]. We wish to leave the universal-time component for the quiet days here and to take up the discussion of the universal-time dependence of magnetic disturbance K in a following paper where the results for disturbed days will be discussed.

Returning to the local-time variation, which is the more definite diurnal change in these quiet days, a matter of interest arose when comparison was made of the 9 yearly average diurnal variations for each of the 9 years separately. These are shown in Figure 3 and indicate a consistent change in form over these years. This time-span almost covers a sunspot cycle and might suggest that the local time diurnal variation changed with this cycle. The character of disturbance in general probably changes appreciably over this interval. In the year 1947 there was much recurrent activity, whereas in 1947 near sunspot maximum the activity may have been largely confined to nonrecurrent magnetic storms.

The principal change pictured in Figure 3 roughly from sunspot minimum (1944.2) to maximum (1947.5). Possibly there was a change over these years in the relative amounts of disturbance having different forms of local time diurnal variation. Disturbance in general appears to have been about as large in 1943 as in 1947. The yearly average K number for 1943 for the 10 quiet days for these six observatories was 1.28, essentially the same as that for 1947, which was 1.26. (The yearly mean K 's for these two years were 2.45 and 2.44, respectively, while the corresponding yearly mean magnetic character figures, C , were 0.71 and 0.69 [Hofmann and Weisman, 1949]).

As to the origin of the diurnal variation shown in Figures 1 and 3, it might be thought, in view of the recognized difficulties in scaling the widest ranges of K , that this lies principally in the inclusion, as K material, of some of the quiet-day daily variation of the field, S_a [Bartels, Eck, and Johnston, 1939, see p. 413, sec. 4]. This would probably produce a maximum near midday, accentuated near the summer solstice [Intosh, 1959, see p. 535], when S_a is relatively large. Figure 1 shows such a maximum. (Since five of the six observatories used are in the northern hemisphere, the June solstice may be considered the summer solstice.) The minimum, however, should probably be flatter than the maximum and occur during the night. Figure 1 does not show such a flat minimum, but a rather sharp one toward the end of the night.

Further, since the amplitude of the daily variation of the field is greater at sunspot maximum than at sunspot minimum, the difficulty in including all S_a variability from K may be somewhat greater near sunspot maximum even though the attempt is made to take account of this in the scaling. Thus a midday maximum from this source might be expected to be more prominent in years of more sunspots, which would be in accord with Figure 3. This variation of the amplitude of the quiet-day daily variation of the field with the solar cycle is well illustrated in the Greenwich Comparative Sunspot and Geomagnetic (S_a range) Curves 1841-1950 [Royal Greenwich Observatory, 1955]. These indicate that the ranges of the quiet-day diurnal variation of both H and D at Greenwich were roughly 50 per cent greater in 1947 than in 1943. Nevertheless, we feel that the major part of the change shown in Figure 3 probably involves disturbance. In view of the care taken by those who scale K to avoid including as K material effects arising from the variability of S_a , the amplitude of the midday maximum in 1947 seems relatively too large to have come mainly from that source. Moreover, the sharp minimum of 3 hours is difficult to explain on the basis of S_a , because the variation of the field on quiet days is relatively small all through the night. The variation for 1943 has a fairly large amplitude, but the principal maximum in that year is near midnight, which could hardly come from the variability of the quiet-day daily variation of the field. This suggests that in 1943 the larger

part of the daily variation of these K numbers for the quiet days arises from disturbance of the kind more prevalent on disturbed days. In support of this, the form of the variation for 1943 (that for 1942 and that for 1944 are roughly similar) resembles somewhat that for all days [Nicholson and Wulf, 1955] and for the more disturbed days.

In our sample study of the 5 quiet days the averages of the local-time variation for the 6 months, as representative of the whole year, were also obtained for the years 1943 and 1947 separately. The result for 1947 resembled closely that for the 10 quiet days shown in Figure 3, suggesting that the disturbance in 1947 was fairly homogeneous over the 10 quiet days. For 1943, however, the 5 quiet days yielded a result differing appreciably in form from that for the 10 quiet days shown in Figure 3, tending toward that for 1947, and suggesting that some disturbance having a midday maximum, as in 1947, was present also in 1943. Magnetic disturbance having a local-time variation with a daytime maximum has been recognized at high latitudes [Mayaud, 1956]. Possibly Figure 3 is evidence that such disturbance may also occur at lower latitudes, especially near sunspot maximum.

These points suggest that in the 10 quiet days per month of these years, the relative amounts of disturbance having different forms of diurnal variation changed in a major way over these years, possibly with the solar cycle.

From our studies of the local-time component for all days and for the more disturbed days we are also inclined to believe that the major part of the variation for the quiet days, including 1947, is due to disturbance, since there are suggestions of a behavior in the more disturbed days somewhat similar to that in the quiet days.

If this tentative inference is correct, it suggests that a factor affecting the diurnal variation of disturbance may have changed over these years. We feel that a part of the diurnal variation of the irregular geomagnetic fluctuations in general might arise from disturbance produced in dynamo action [Nicholson and Wulf, 1958], in the presence of a large-scale atmospheric circulation in the ionosphere. A change with the solar cycle in the large-scale wind system in the ionosphere seems possible, and might affect the diurnal variation of disturbance.

Acknowledgment. The assistance of Miss Myrtle Richmond who helped with the computational work is gratefully acknowledged.

REFERENCES

- Bartels, J., N. H. Heck, and H. F. Johnston, The three-hour-range index measuring geomagnetic activity, *Terrestrial Magnetism and Atmospheric Elec.*, *44*, 411-454, 1939.
- Bartels, J., and H. F. Johnston, Main features of the daily magnetic variations at Sitka, Cheltenham, Tucson, San Juan, Honolulu, Huancayo and Watheroo, *Terrestrial Magnetism and Atmospheric Elec.*, *44*, 455-469, 1939.
- Howe, H. H., and E. K. Weisman, Geomagnetic indices *C* and *K*, 1948, Bull. No. 12b, Association of Terrestrial Magnetism and Electricity, International Union Geodesy and Geophysics, Washington, D. C., 1949.
- Johnston, H. F., W. E. Scott, and E. Balsam, Geomagnetic indices *C* and *K*, 1940-1946, Bull. No. 12, 1947, Bull. No. 12a, Association of Terrestrial Magnetism and Electricity, International Union Geodesy and Geophysics, Washington, D. C. 1948.
- Mayaud, P. M., Rapports scientifiques des Expéditions Polaires Françaises S IV 2, *Ann. géophys.* *12*, 84-101, 1956.
- McIntosh, D. H., On the annual variation of magnetic disturbance, *Phil. Trans. Roy. Soc. London A*, *251*, 525-552, 1959.
- Nicholson, S. B., and O. R. Wulf, The diurnal variation of irregular geomagnetic fluctuations, *Geophys. Research*, *60*, 389-394, 1955.
- Nicholson, S. B., and O. R. Wulf, The diurnal variation of irregular geomagnetic fluctuations, I *J. Geophys. Research*, *63*, 803-806, 1958.
- Royal Greenwich Observatory, Sunspot and geomagnetic-storm data, Her Majesty's Stationery Office, London, 1955.

(Manuscript received January 7, 1961.)

Energy Transformation and Vertical Flux Processes over the Northern Hemisphere¹

CLAYTON E. JENSEN

*Air Force Institute of Technology
Wright-Patterson Air Force Base, Ohio*

Abstract. Three integrals in the mechanical energy equation have been statistically evaluated as a function of a broad division of eddies in the atmosphere over the northern hemisphere for January and April 1958. The integral of the product of the instantaneous pressure change and the specific volume exhibits two significant positive (potential to kinetic energy) modes, one in the boundary region from 1000 to 850 millibars and the other in the middle troposphere from 700 to 500 millibars. The integral values of the product of the instantaneous pressure change and the potential energy are about an order of magnitude larger than the values computed for the integral representing the transport of kinetic energy. The three integral values were introduced into the mechanical energy equation and the residual was taken, ideally at least, to represent the dissipation rate for the horizontal kinetic energy within various regions of the atmosphere up to 50 millibars. These dissipation rates compare favorably with the appropriate estimates by Brunt.

Introduction. As was first suggested by Burgules [1903], the significant action in the transformation process which produces kinetic energy in the atmosphere appears to be the simultaneous rising of warm air and sinking of cold air. One might suspect that the more vigorous energy transformation processes take place in the lower and middle troposphere where baroclinicity is more pronounced and where kinetic energy is produced at the expense of potential energy. In the upper troposphere and lower stratosphere there is some evidence [White and Nolan, 1959] of a reversal of this process; i.e., kinetic energy is transformed into potential and internal energy. Studies such as the one first cited and others by White and Saltzman [1956], Saltzman and Fleisher [1959], and Winnielsen [1959] have had the limitations of a small spatial sample and/or the use of values in the vertical motion term for only one surface or for only one layer.

In this paper we have evaluated the magnitude of the energy transformation processes on a hemispheric basis as a function of height and also as a function of a rather broad division of eddies, i.e., the so-called standing eddies and

the transient eddies. In addition we have evaluated certain vertical flux processes involving potential and kinetic energy for regions of the atmosphere bounded by closed pressure surfaces. These measures of the vertical flux of energy have been introduced into the energy balance equation along with or as part of the values for the energy transformation process for the region in question. Any residual may be taken, ideally at least, as an approximation to the magnitude of the dissipative action within the same region.

The mechanical energy equation. To obtain the mathematical expression which represents the balance equation for mechanical energy, we follow a technique introduced by Starr [1951] with x, y, z, t coordinates and later adapted by Phillips [1954] to the x, y, p, t coordinate system. The horizontal vector equation of motion per unit mass, with hydrostatic equilibrium assumed and pressure used as the vertical coordinate, is

$$\frac{d\mathbf{V}}{dt} + f\mathbf{k} \times \mathbf{V} + \nabla_p \varphi + \mathbf{F} = 0 \quad (1)$$

where \mathbf{V} is the horizontal wind vector, f is the horizontal component of the Coriolis force, \mathbf{k} is the unit vertical vector, $\varphi = gz$ is the geopotential of an isobaric surface, ∇_p is the horizontal 'del' operator along a constant pressure surface, and \mathbf{F} is the horizontal friction force.

If we take the scalar product of (1) and \mathbf{V}

¹ This research was performed at the Massachusetts Institute of Technology in partial fulfillment of the requirements for the degree of Doctor of Philosophy. The author's current address is Headquarters 3rd Weather Wing, Offutt Air Force Base, Nebraska.

and expand the total derivative, letting $\mathbf{V}^2/2 = K$ be the horizontal kinetic energy, $\varphi = gz = \Phi$ be the potential energy, and $\mathbf{V} \cdot \mathbf{F} = D$ be the rate of dissipation of kinetic energy, all per unit mass, we obtain

$$\frac{\partial K}{\partial t} + \mathbf{V} \cdot \nabla_p K + \frac{dp}{dt} \frac{\partial K}{\partial p} + \mathbf{V} \cdot \nabla_p \Phi + D = 0 \quad (2)$$

If the hydrostatic equation in the form

$$\partial \omega / \partial p + \alpha = 0 \quad (3)$$

where $\alpha = 1/\rho$ is the specific volume, and the continuity equation

$$\partial \omega / \partial p + \nabla_p \mathbf{V} = 0 \quad (4)$$

where $\omega = dp/dt$ are used (ω is the instantaneous pressure change representing vertical motion in the coordinate system where pressure replaces height as the vertical axis), we may write (2), the mechanical energy equation for horizontal motion, as

$$\frac{\partial K}{\partial t} + \nabla_p \cdot (K + \Phi) \mathbf{V} + \frac{\partial(\omega K)}{\partial p} + \frac{\partial(\omega \Phi)}{\partial p} + \omega \alpha + D = 0 \quad (5)$$

Integrating (5) over the mass of the atmosphere between closed constant pressure surfaces, we obtain

$$\begin{aligned} & \frac{\partial}{\partial t} \iiint K \frac{dp}{g} dx dy \\ & + \iint [(\omega K)_u - (\omega K)_l] \frac{dx}{g} dy \\ & + \iint [(\omega \Phi)_u - (\omega \Phi)_l] \frac{dx}{g} dy \\ & + \iiint \omega \alpha \frac{dp}{g} dx dy \\ & + \iiint D \frac{dp}{g} dx dy = 0 \end{aligned} \quad (6)$$

where the subscripts u and l refer to the upper (1000-mb) and lower (50-mb) boundary values, respectively.

Using the definition of the 'source' of kinetic energy given by Starr [1951] as the integral of

the scalar product of the horizontal wind and the gradient of pressure in the x, y, z, t coordinate system, this integral may be transformed into the x, y, p, t coordinate system as follows:

$$\begin{aligned} & \iiint \mathbf{V} \cdot \nabla_h p \, dx \, dy \, dz \\ & = \iint [(\omega \Phi)_u - (\omega \Phi)_l] \frac{dx}{g} dy \\ & + \iiint \omega \alpha \frac{dp}{g} dx dy \end{aligned} \quad (7)$$

where the integration is taken between two closed pressure surfaces.

Thus, we define the following processes which are represented in (6):

$$\begin{aligned} & \iint [(\omega \Phi)_u - (\omega \Phi)_l] \frac{dx}{g} dy \\ & + \iiint \omega \alpha \frac{dp}{g} dx dy \\ & = \text{transformation process} \\ & \iint [(\omega K)_u - (\omega K)_l] \frac{dx}{g} dy \\ & = \text{transport process} \end{aligned}$$

$$\iiint D \frac{dp}{g} dx dy = \text{dissipation process}$$

where, of course, the sum of the three processes make up the observed change in the horizontal kinetic energy of the portion of the atmosphere being sampled between two closed pressure surfaces. It is noteworthy that the integral involving the boundary values of $\omega \Phi$ represents in essence transport of potential energy, and yet it more rigorously belongs with the energy transformation process and is a consequence of the slope of pressure surfaces relative to constant heights.

The vertical motion equation. In the present study an attempt is made to examine the integral of $\omega \alpha$ as well as the integrals of $\omega \Phi$ and ωK horizontally over the northern hemisphere north of 20°N and vertically by pressure increment within the layer from 1000 to 50 mb.

The crucial measurement is the one involving the instantaneous pressure change or the vertical motion term, ω . If we expand the total derivative of temperature with respect to time in tl

ation which expresses the first law of thermodynamics for our atmosphere, and assume adiabatic changes, we obtain

$$-\frac{\omega}{\rho g} = -\frac{\partial T/\partial t + \mathbf{V} \cdot \nabla_p T}{g/c_p + \partial T/\partial z} \quad (8)$$

where T is the temperature and c_p is the specific heat at constant pressure.

It is interesting to note that the left-hand side of (8) is the well-known approximate expression for the vertical velocity, w [see Charney, 1948, for example]. The expression on the right-hand side of (8) was evaluated for several pressure layers from actual upper-air data taken every 6 hours at approximately 100 stations over the northern hemisphere for the months of January and April 1958. All the individual components of the right-hand side of (8) except one are readily determined from available constant-pressure data. The one exception is the term $\mathbf{V} \cdot \nabla_p T$, which represents the mean temperature advection within the layer. The evaluation of this term was accomplished in a manner somewhat similar to the technique first presented by Panofsky [1944].

The final expression submitted in the finite-difference form for machine computational purposes and designed to give the 12-hour time-averaged vertical velocity (with equivalence to (8)) for a specified isobaric layer and for a specified station, using the data from that station only, is as follows:

$$\begin{aligned} -\frac{\omega}{\rho g} = & -\frac{\frac{1}{2}(T_1 + T_2) \big|_{t=0} - \frac{1}{2}(T_1 + T_2) \big|_{t=-12}}{\frac{g}{c_p} + \frac{1}{2} \left[\frac{T_2 - T_1}{\Delta h} \bigg|_{t=0} + \frac{T_2 - T_1}{\Delta h} \bigg|_{t=-12} \right]} \\ & + \frac{\left[V_1 V_2 \frac{\Delta \sigma}{\Delta h} \bigg|_{t=0} + V_1 V_2 \frac{\Delta \sigma}{\Delta h} \bigg|_{t=-12} \right] \frac{1}{4} [(T_1 + T_2) \big|_{t=0} + (T_1 + T_2) \big|_{t=-12}] \frac{\pi f}{(980)(180)}}{\frac{g}{c_p} + \frac{1}{2} \left[\frac{T_2 - T_1}{\Delta h} \bigg|_{t=0} + \frac{T_2 - T_1}{\Delta h} \bigg|_{t=-12} \right]} \quad (9) \end{aligned}$$

where the subscripts 1 and 2 refer to the higher and the lower pressure surfaces, respectively, $\Delta \sigma$ is the change in wind direction in degrees of the compass computed as the wind direction at the lower pressure surface minus the wind direction at the higher pressure surface measured positive clockwise from north, Δh is the thickness between two pressure surfaces, and v is the wind speed.

Statistical analysis. Computation of the vertical motion by means of (9) permits examination of the variability of this quantity itself and also of its statistical interrelationships with parameters such as α , Φ , and κ as a function of specified layers up to the top of the atmosphere. Seven layers were selected for this study, namely, 1000-850, 850-700, 700-500, 500-300, 300-200, 200-100, and 100-50 mb. The vertical-motion data computed by (9) and the associated 12-hour mean temperatures were tabulated by 12-hour periods for the months of January and April 1958 and for approximately 100 stations located north of 20°N. The mean temperature was used in lieu of the specific volume in the actual evaluation of that portion of the transformation process represented by the integral of $\omega \alpha$.

In the statistical treatment of the integrals containing $\omega \alpha$, $\omega \Phi$, and $\omega \kappa$, we follow closely the technique first applied to hemispheric data by Starr and White [1952] and adopt the bar notation for the time average of a quantity between specified limits, the bracket for the zonal average, and the brace for the meridional average. Further, the notation of a single prime denotes the eddy or turbulent fluctuations of the quantity about its time average, a double prime the deviation about its zonal average, and a triple prime the deviation about its meridional average.

Inasmuch as we do not integrate over the entire atmosphere nor do we perhaps take a long enough sample in time, it is possible for

the mean-motion terms, such as $\{\bar{\omega}\} \{\bar{T}\}$, to have questionable nonzero values, which, however, we cannot hope to measure separately. Thus, to get a more representative value of the total contribution from the individual transformation and transport integrals, we subtract the possibly spurious mean-motion term from both sides of the eddy correlation equations,

TABLE 1. Mean Vertical Motion in Centimeters per Second by Pressure Layer and by Latitude for January 1958

Pressure Layer, mb	Latitude						
	20°	30°	40°	50°	60°	70°	80°
1000-850	0.281	0.728	1.361	1.392	0.678	0.567	0.347
850-700	-0.036	-0.128	0.317	0.711	0.469	0.186	0.011
700-500	-0.067	-0.214	0.125	0.503	0.086	-0.206	0.133
500-300	0.069	-0.061	0.008	0.411	0.197	-0.253	-0.044
300-200	-0.531	0.403	0.033	0.194	-0.117	-0.642	-0.306
200-100	-0.036	0.008	-0.081	-0.061	-0.106	-0.039	-0.081
100-50	0.007	-0.017	0.031	0.014	0.039	0.208	-0.139

with the result (shown only for the case of ω and T) that

$$\frac{1}{g} \{ \{(\overline{\omega K})\} \} - \frac{1}{g} \{ \{(\overline{\omega})\} \} \{ \{(\overline{K})\} \}$$
$$\{ \{ \overline{\omega T} \} \} - \{ \{ \overline{\omega} \} \} \{ \{ \overline{T} \} \} = \{ \{ \overline{\omega' T'} \} \}$$
$$+ \{ \{ \overline{\omega'' T''} \} \} + \{ \{ \overline{\omega''' T'''} \} \} \quad (10)$$
$$= + \iint \omega K \frac{dx}{g} dy \quad (11)$$

The following equivalences are noted relative to the mechanical energy equation (6), and we now introduce the vertical integration with respect to pressure as denoted by the parentheses and further assume that lower (50-mb) boundary values vanish in the integrals of $\omega\Phi$ and $\omega\kappa$:

$$\frac{R}{pg} \{ \{(\overline{\omega T})\} \} - \frac{R}{pg} \{ \{(\overline{\omega})\} \} \{ \{(\overline{T})\} \}$$
$$= + \iiint \omega \kappa \frac{dp}{g} dx dy \quad (11)$$
$$\frac{1}{g} \{ \{(\overline{\omega \Phi})\} \} - \frac{1}{g} \{ \{(\overline{\omega})\} \} \{ \{(\overline{\Phi})\} \}$$
$$= + \iint \omega \Phi \frac{dx}{g} dy \quad (12)$$

It is seen from (10) that the integrals of $\omega\Phi$, and $\omega\kappa$ can each be resolved into contributions from three eddy processes, namely, the transient eddies, the standing eddies (zonal), and the standing eddies (meridional).

Vertical motion and temperature results. The January latitudinal averages of the vertical motion, in centimeters per second, for each of the pressure layers are listed in Table 1, and the corresponding values for April are listed in Table 2.

The January and April mean values of temperature in degrees Kelvin as a function of pressure layer and latitude are listed in Tables 3 and 4.

Figure 1 shows the January zonal averages of the atmospheric vertical motion as a function of latitude.

TABLE 2. Mean Vertical Motion in Centimeters per Second by Pressure Layer and by Latitude for April 1958

Pressure Layer, mb	Latitude						
	20°	30°	40°	50°	60°	70°	80°
1000-850	0.152	0.225	0.614	0.861	1.133	0.639	0.05
850-700	-0.063	0.175	0.517	0.678	0.697	0.428	0.12
700-500	0.192	0.117	0.042	0.131	0.161	-0.022	-0.01
500-300	0.003	-0.311	-0.722	-0.320	0.631	-0.064	-0.28
300-200	0.222	0.028	-0.256	0.372	0.375	0.475	0.51
200-100	-0.065	-0.083	-0.072	0.015	-0.010	0.014	-0.10
100-50	-0.041	0.026	0.094	0.066	0.024	0.007	0.01

TABLE 3. Mean Temperature in Degrees Kelvin by Pressure Layer and Latitude for January 1958

Pressure Layer, mb	Latitude						
	20°	30°	40°	50°	60°	70°	80°
1000-850	288.8	282.2	272.6	264.7	258.4	251.9	247.6
850-700	283.0	276.1	268.6	262.4	256.8	251.0	247.9
700-500	273.0	265.6	257.8	251.6	246.6	242.6	240.3
500-300	250.9	244.7	238.3	232.7	228.3	224.8	223.8
300-200	228.3	224.6	221.3	218.9	216.3	212.8	210.9
200-100	207.8	212.0	216.8	216.9	214.6	212.9	212.0
100-50	204.0	209.7	215.8	218.6	217.5	215.0	213.4

latitude and as a function of the following layers: 1000-50, 850-50, 700-50, and 500-50 mb. Figure 2 is the comparable presentation for April. These figures also have the zonal average of temperature for the layer from 1000 to 50 mb plotted as a function of latitude. In general, the combined effects of the vertical motion and temperature curves in Figures 1 and 2 tend to support the historical feature of a strong indirect meridional cell in mid-latitudes with the suggestion of less intense direct cells to the north and to the south. The indirect cell, or one with rising cold air and sinking warm air, appears to be quite striking and gives credence to ideas presented by *Ferrel* [1856] and subsequently modified by *Rossby* [1941]. The peak value of the cellular motion of approximately 0.3 cm/sec for the layer from 700 to 50 mb compares reasonably well with the peak value of about 0.1 cm/sec found for the month of February 1959 by *Saltzman and Fisher* [1959] who used vertical-motion data obtained by the numerical weather prediction method. *Phillips* [1954] found that the simple

baroclinic waves in his two-level, quasi-geostrophic model are accompanied by a weak meridional circulation pattern, similar to the observational evidence of Figures 1 and 2 but with a somewhat reduced maximum vertical motion of about 0.05 cm/sec.

In the long-time mean, the area under each of the vertical-motion curves in Figures 1 and 2 should be zero when closed pressure surfaces over the whole hemisphere are considered. The actual values of the integrals are listed in Table 5.

The integrated values of the vertical motion given in Table 5 decrease in algebraic magnitude as the lower layers of the atmosphere are consecutively eliminated from the computations. In fact, the values for April pass through the null point. A zero value for the integral appears to occur for a lower boundary of about 700 mb when the combined January and April data (Table 5) are considered. This suggests the influence in the lower atmosphere of two effects which were neglected in the computations. First, the nonadiabatic effects would have to act in

TABLE 4. Mean Temperature in Degrees Kelvin by Pressure Layer and by Latitude for April 1958

Pressure Layer, mb	Latitude						
	20°	30°	40°	50°	60°	70°	80°
1000-850	291.5	285.9	279.3	273.3	267.5	257.1	248.5
850-700	285.7	281.4	274.3	267.3	261.4	254.2	248.8
700-500	274.5	269.3	262.5	255.6	250.0	243.9	238.9
500-300	252.6	247.3	241.4	235.5	231.5	228.5	226.5
300-200	228.8	224.7	222.3	220.7	219.7	222.3	225.3
200-100	208.4	210.6	214.8	219.1	221.7	223.6	224.8
100-50	205.2	210.7	216.2	220.0	222.0	223.0	221.6

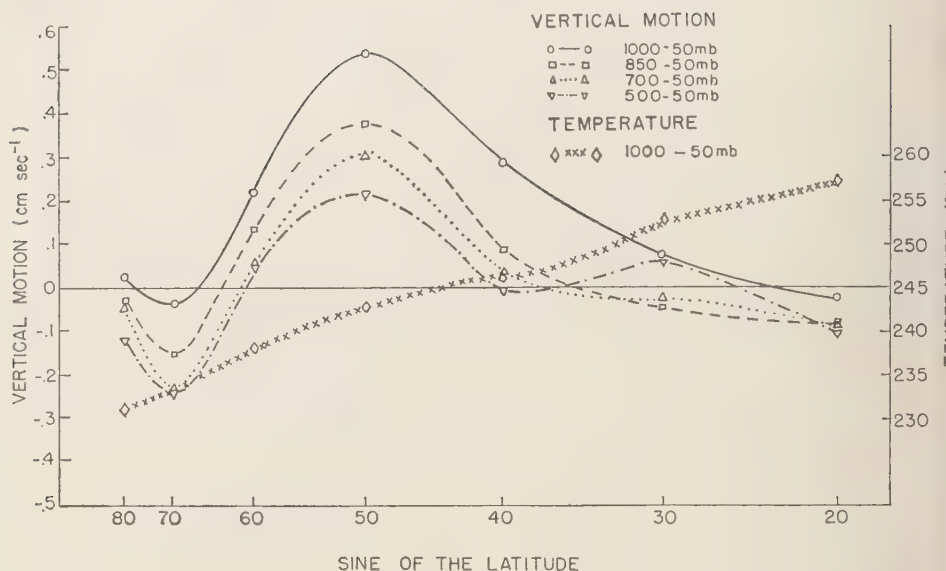


Fig. 1. Mean vertical motion and mean temperature as a function of latitude and pressure layer for January 1958.

the sense of a net cooling in order to influence the computed vertical velocities in the upward or positive sense. For the lower layers of the atmosphere, at least, this would seem to rule out the neglect of the nonadiabatic effects as being the dominant cause of the seemingly large positive values of vertical motion. The second influence involves the nongeostrophic effects within the friction layer which may be defined as the portion of the atmosphere below 800 meters. Within this layer, the wind veers with height approximately in accordance with the *Ekman* [1902] spiral, where the angle between the actual wind at anemometer level and the geostrophic wind at the top of the friction layer is about 20° [Taylor, 1915]. The veering of the

wind with height implies warm-air advection from the thermal wind relation. This latter relation was applied in evaluating the temperature advection term in (9), but actual rather than geostrophic winds were used; thus there is a strong possibility that nongeostrophic effects produced a systematic positive error in the computations of vertical motion for the layer from 1000 to 850 mb. Some of this effect undoubtedly extends into the next higher layer (850 to 700 mb) since there is an indication in Tables 1 and 2 that the vertical motion for the lower layer is perhaps too great on the positive side. In general, Table 5 shows that the layer from 700 to 50 mb is fairly representative of quasi-geostrophic conditions in the free atmosphere when the integrated values of vertical motion for both January and April are considered.

TABLE 5. Integrated Values of the Vertical Motion in Centimeters per Second by Pressure Layer between 20°N and 80°N for January and April 1958

Pressure Layer, mb	January	April
1000-50	+0.1702	+0.1330
850-50	+0.0501	+0.0626
700-50	+0.0149	-0.0016
500-50	+0.0045	-0.0518

Energy transformation due to vertical motion and temperature covariance. The energy transformation rates due to the transient eddies and to the standing eddies (zonal) shown in (10) have been added, and the results appear in tabular form in Table 6 and as a function of pressure-height in Figure 3. The curves for January and April show the same two positive modes, the lower one in the boundary region from 1000 to 850 mb and the other in the mid-tropospheric region from

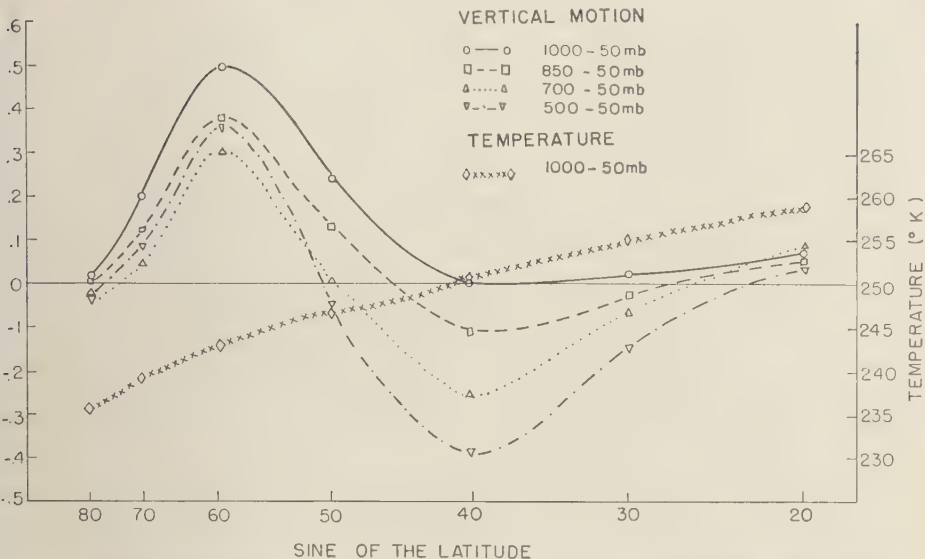


Fig. 2. Mean vertical motion and mean temperature as a function of latitude and pressure layer for April 1958.

to 500 mb. The lower and more predominant side in the boundary region is questionable because of previously discussed nongeostrophic and nonadiabatic effects. On the other hand, this could be a realistic phenomenon and a reflection of observed low-level convergence and divergence. The prominent positive mode in the lower-tropospheric region centered at about 600 mb is quite significant and not wholly unexpected. There seems to be little doubt that this is a feature in the atmosphere where energy trans-

formation processes take place in a very effective, organized manner. The curve for April gradually turns negative at about 200 mb and remains so above that surface. The curve for January oscillates between negative and positive values above 400 mb and finally turns to negative values above 100 mb.

Saltzman and Fleisher [1959] obtained a value of 2.682 ergs/cm² sec mb for the total or net rate of energy transformation for the month of February 1959. They used thickness data for

TABLE 6. Energy Transformation Rates Due to the Transient Eddies and to the Standing Eddies (Zonal) Listed as a Function of Pressure Layer for January and April 1958 (Units are ergs/cm² sec mb. Positive values signify a transformation from potential to kinetic energy.)

Pressure Layer, mb	January			April		
	Transient Eddies	Standing Eddies (zonal)	Total	Transient Eddies	Standing Eddies (zonal)	Total
1000-850	3.068	5.905	8.973	0.616	5.845	6.461
850-700	2.338	1.739	4.077	2.410	1.202	3.612
700-500	3.495	3.531	7.026	1.953	2.238	4.191
500-300	1.057	3.233	4.290	2.146	0.219	2.365
300-200	-1.392	1.057	-0.335	1.450	-0.950	0.500
200-100	-0.246	1.624	1.378	-0.594	-0.273	-0.867
100-50	-0.422	-0.405	-0.827	-0.144	-0.625	-0.769

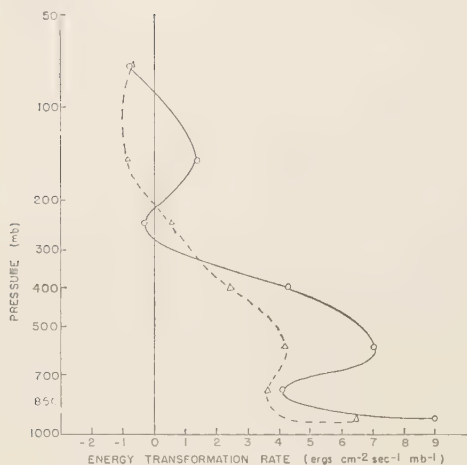


Fig. 3. Energy transformation rates due to the combined effect of the transient eddies and the standing eddies (zonal) as a function of pressure-height for the layer from 1000 to 50 mb for January 1958 (solid curve) and April 1958 (dashed curve).

the layer from 850 to 500 mb to obtain the field of temperature with values of ω computed at 600 mb obtained from the Joint Numerical Weather Prediction Unit two-parameter, quasi-geostrophic, adiabatic, frictionless model. Values of the transformation rates between 850 and 500 mb for January (Table 6), integrated and combined with the tropospheric contribution of -1.488 ergs/cm² sec mb for the remaining standing eddies (meridional), would yield a net transformation rate of 4.274 ergs/cm² sec mb, which is somewhat larger than the corresponding one obtained by Saltzman and Fleisher.

Energy transformation due to vertical motion and potential energy covariance. From (6) and (7) it is shown that the integral of $\omega\alpha$ plus the integral of $\omega\Phi$ jointly represent the energy transformation process. The integral of $\omega\Phi$ is evaluated, however, as a boundary value of the flux of potential energy through some arbitrary pressure surface with units of ergs/cm² sec.

Since the integrals of $\omega\Phi$ and $\omega\kappa$ in (6) both entail boundary values, it would appear reasonable to dispense with both integrals by specifying boundary conditions such that both integrals vanish. This seems to be not too unrealistic, since $\omega \rightarrow 0$ at the outer limits of the atmosphere and since it is difficult to visualize a flux of kinetic or potential energy through the surface

of the earth. If this is the case, the roles played by the integrals of $\omega\Phi$ and $\omega\kappa$ in (6) are probably nothing more than those involving some internal adjustments which are required in response to the energy transformation function represented by the integral of $\omega\alpha$.

It appears worthwhile, however, in consideration of the energy balance equation, to investigate the magnitude of the flux of kinetic energy as well as that of the flux of potential energy. The former is treated in detail in a later section, and we are indebted to Roberts [1960] for the values of the flux of potential energy which are listed in Table 7 as a function of pressure layer and eddy process for January 1958.

The contribution from the transient eddies in Table 7 is negative up to 300 mb and becomes positive within the layers from 300 to 200 mb and 100 to 50 mb. This suggests that rising motion is associated with troughs and sinking motion with ridges in the troposphere, whereas the reversal of the sign of the flux above 300 mb suggests an opposite-phase relationship in certain layers of the stratosphere, with ascending motion in ridges and descending motion in troughs. This phase shift has been previously noted for the upper stratosphere by Kochanski [1954] and Austin and Krawitz [1956].

Energy transport due to vertical motion and kinetic energy covariance. This transport process is represented by (13). A cursory examination of the hemispheric distribution of the covariance between vertical motion and horizontal kinetic energy shows that, in general, and for both months, there is a net upward flux of kinetic

TABLE 7. Vertical Flux of Potential Energy Function of Pressure Layer and Eddy Process for January 1958
(Units are ergs/cm² sec. Positive values signify an upward flux of potential energy.)

Pressure Layer, mb	Transient Eddies	Standing Eddies (zonal)	Standing Eddies (meridional)
1000-850	-322.5	-2749.5	-1370.0
850-700	-1263.2	-207.3	-2158.0
700-500	-241.3	+1049.4	-1672.0
500-300	-465.2	+747.2	-570.0
300-200	+756.2	+270.9	+201.0
200-100	-102.8	+338.3	+108.0
100-50	+107.9	-48.0	-78.0

TABLE 8. Vertical Transport of Kinetic Energy Due to the Three Eddy Processes Integrated with Respect to Longitude and Listed as a Function of Pressure Layer for January and April 1958 (Units are ergs/cm² sec. Positive values signify an upward transport of horizontal kinetic energy.)

Pressure Layer, mb	January			April		
	Transient Eddies	Standing Eddies (zonal)	Standing Eddies (meridional)	Transient Eddies	Standing Eddies (zonal)	Standing Eddies (meridional)
1000-850	306.43	80.37	70.79	219.79	156.18	23.26
850-700	159.35	-22.83	34.40	220.99	91.92	45.67
700-500	18.40	-158.08	5.12	71.84	9.20	-6.72
1000-500	147.06	-37.37	46.71	159.95	78.15	19.71

gy due to the transient eddies for the layers 1000-850, 850-700, and 700-500 mb. This seems contradictory to intuitive reasoning, which would lead one to suspect that the flux of kinetic energy should be directed downward in order to balance the dissipation of kinetic energy in the boundary layer. By the same token, however, one might reason that there should be an upward flux of kinetic energy in order to help maintain the westerlies against dissipative action. It has been shown by *Starr and White* [1951], for example, that angular momentum is transported downward mainly by the transient eddies and that the flux of momentum also helps to maintain the westerlies. Both the vertical transportation of kinetic energy and the poleward transportation of angular momentum are directed against the gradient of a wind velocity. The significance of a positive value of the covariance between vertical motion and horizontal kinetic energy is that rising motion must occur predominantly in regions of maximum wind velocity. In the dynamic sense, this means that on the average there must be stronger winds to the lee of regions where rising motion is normally evident. The vertical transport of kinetic energy due to the three eddy processes is shown in Table 8. *Energy balance results.* From a scale factor analysis the values obtained for the integral of kinetic energy should be about an order of magnitude larger than the values for $\omega\kappa$, or approximately equal to the ratio of the Coriolis parameter to the eddy vorticity. This appears to be borne out very well, as can readily be verified by comparing Tables 7 and 8.

The integrals in the mechanical energy equation (6) representing energy transformation and flux processes have been individually

evaluated for various pressure layers. The sum of these integral values should balance, ideally at least, the dissipation of horizontal kinetic energy within the region of the atmosphere being sampled. In this analysis, the contributions due to the standing eddies (meridional) will be omitted, since these values are not considered to be as stable as the contributions computed for the two other eddy processes. Using Tables 6, 7, and 8 for data for the integral values of $\omega\alpha$, $\omega\Phi$, and $\omega\kappa$, respectively, we have computed energy dissipation rates for selected layers of the atmosphere for January 1958 (Table 9). For the layers which extend to 50 mb the flux of energy across that pressure surface is neglected. This appears to be justified by the small net flux of potential energy shown in Table 7 for the layer from 100 to 50 mb. In addition, the flux of kinetic energy is probably negligible at high levels; for example, the kinetic energy flux has been computed by *Hansrote and Lambert* [1960] to be approximately 10 ergs/cm² sec for the layer from 200 to 100 mb for April 1958. It is also assumed (Table 9) that the transports of potential energy and kinetic energy vanish at 1000 mb, which is taken to be the surface value of pressure at $Z_0 = 0$. Furthermore, the flux of kinetic energy is neglected above 500 mb, and this seems justified since it is an order of magnitude smaller than the flux of potential energy.

A separation of data in Table 9 is made in order to distinguish between the dissipation rates computed, in general, for rather deep layers, each extending to 50 mb, versus the rates computed for rather shallow layers with depths of from 50 to 75 mb.

Brunt [1941] estimated the energy dissipation rate to be approximately -3000 ergs/cm² sec in

TABLE 9. Energy Dissipation Rates within Certain Atmospheric Layers for January 1958 (in ergs/cm² sec)

Pressure Layer, mb	$\omega\alpha$	$(\omega\Phi)_u$ $-(\omega\Phi)_l$	$(\omega K)_u$ $-(\omega K)_l$	Energy Dissipation Rates
1000-50	+4283.6	0	0	-4283.6
925-50	+3610.7	-3072.0	+386.8	-925.5
850-50	+2937.7	-2271.3	+261.7	-928.1
775-50	+2631.9	-1470.5	+136.5	-1297.9
700-50	+2326.1	-662.4	-1.6	-1662.1
600-50	+1623.5	+808.1	-139.7	-2291.9
500-50	+920.9	+595.1		-1516.0
400-50	+491.9	+282.0		-773.9
300-50	+62.9	+654.6		-717.5
250-50	+79.6	+1027.1		-1106.7
200-50	+96.4	+631.3		-727.7
150-50	+27.5	+235.5		-263.0
100-50	-41.4	+147.7		-106.3
<hr/>				
1000-925	+672.9	+3072.0	-386.8	-3358.1
925-850	+672.9	-800.7	+125.1	+2.7
850-775	+305.8	-800.8	+125.2	+398.8
775-700	+305.8	-808.1	+134.9	+367.4
700-600	+702.6	-1470.5	+138.1	+629.8
600-500	+702.6	+213.0	-139.7	-775.9
500-400	+429.0	+313.1		-742.1
400-300	+429.0	-372.6		-56.4
300-250	-16.8	-372.5		+389.3
250-200	-16.8	+395.8		-379.0
200-150	+78.9	+395.8		-474.7
150-100	+68.9	+87.8		-156.7
100-50	-41.4	+147.7		-106.3

the layer from the ground up to about 1 km, and this compares favorably with the value of -3358.1 ergs/cm² sec in Table 9 for the layer from 1000 to 925 mb. Brunt further estimated the energy dissipation rate to be approximately -2000 ergs/cm² sec in the atmosphere above 1 km, with a total dissipation rate of about -5000 ergs/cm² sec. This latter estimate would correspond to the value of -4283.6 ergs/cm² sec (Table 9) computed for the layer from 1000 to 50 mb.

The curious feature in Table 9 is the appearance of a group of positive values of the dissipation rate within the four shallow pressure layers located between 925 and 600 mb. Also, there is one additional positive value in the layer from 300 to 250 mb. The locations of these positive dissipation rates suggest that on the average they occur just above the friction layer and near

the tropopause, respectively. Both these regions are characterized by wind maxima, although the lower-level jet is certainly not as pronounced as the more familiar upper-level jet. Further, both these regions contain turbulent eddies of the type often referred to as clear air turbulence. It is entirely possible that these eddies are of such scale that they were not detected in the data used in this study and further that these eddies act to produce an amount of kinetic energy sufficient to change the sign of the dissipation rate in the energy balance equation.

Conclusions. The energy-transformation process represented by the integral of $\omega\alpha$ in the mechanical energy equation (6) exhibits two significant positive (potential to kinetic energy) modes, one in the boundary region from 1000 to 850 mb and one in the middle troposphere from 700 to 500 mb. It is probably not too surprising to find that one of the regions of more intense energy-transformation activity is close to the commonly observed region of nondivergence at 600 mb. The occurrence of very intense energy transformations within the boundary region is questionable in view of certain nongeostrophic and nonadiabatic effects which probably influenced the computations of vertical motion in this region. It is conceivable, however, that the computed rates of energy transformation are not spurious for this boundary region, since the rate of energy dissipation increase rather sharply in layers close to the ground. Reference is made to the work of Davidson and Lettau [1957], for example.

The evidence for meridional cell activity appears quite striking and conforms in general with the usual arrangement of an intense mid-latitude indirect cell with direct cells to the north and to the south. The northward phase shift of the cells from January to April seems to be consistent with synoptic evidence of the northward displacement of the baroclinic storm tracks and the semipermanent pressure systems from winter to summer. Granting the existence of the indirect mid-latitude cell, the inclusion of the implied nonadiabatic effects would have resulted in an intensification of the cell activity.

The flux of potential energy has been determined to be about an order of magnitude larger than the transport of kinetic energy, and this is in agreement with an analysis of these w

processes in which appropriate scale factors were used.

As there is no net change in the kinetic energy of the atmosphere over a long period of time, the values corresponding to the integrals of $\omega\alpha$, $\omega\Phi$, and $\omega\kappa$ were introduced into the mechanical energy equation (6), and the residual was taken to represent the dissipation rate for the horizontal kinetic energy. This was done for various regions of the atmosphere. The resulting values agree very well with other independent estimates, being about -3000 ergs/cm² sec in the friction layer and about -1000 ergs/cm² sec above the friction layer. There are certain shallow layers in the atmosphere, located just above the friction layer and also near the tropopause, where the dissipation rates are computed to be positive. This indicates that the energy-transformation process, as measured, is not producing the required amount of kinetic energy for a proper atmospheric energy balance, and thus the dissipative action must unrealistically produce rather than destroy kinetic energy. It is more likely, of course, that the dissipative action is small in these regions and that there are other eddies, not discerned by the methods of this study, which tend to make up the deficit of kinetic energy in the energy balance. There is the suggestion that eddies representing clear air turbulence, for example, are of the type of process which has not been detected in the present analysis.

This study has been based upon data from two specific months (January and April 1958), and it is not necessarily a representative sample in time. Further, the portion of the atmosphere that was sampled covered only the northern hemisphere horizontally between 20° and 80° and vertically between 1000 to 50 mb, which is not necessarily a representative sample in space. The only data that were tested for significance were the set of station values of $\bar{\omega T}$ for a few pressure layers. The 95 per cent confidence factor of $2\sigma/\sqrt{N}$ was used. The sets of data examined produced confidence limits which indicate a reasonable degree of significance, e.g., a mean value of 207 ± 75 . A more rigorous significance test could be made only with considerable effort; daily maps of vertical motion, temperature, kinetic energy, and potential energy would be needed in order to get daily values of the various covariances which could then be

tested in sets. On the other hand, one could perform the same type of study as the present one, but for several additional months, and then test the significance of the acquired set of monthly mean values of the covariances.

Acknowledgments. I am grateful to Dr. Robert M. White for his suggestion of the topic of this report and for his initial guidance. Much appreciation is due Professor Victor P. Starr for his never-ending encouragement and indispensable advice throughout the 14 months devoted to this study.

Basic meteorological data were machine-processed in an expert manner by the staff of the Air Weather Service Climatic Center, and these processed data formed the foundation for this study. In particular, I wish to acknowledge the cooperation of Lt. Col. George W. Moxon, the Director of the Center.

I am also indebted to the staff of the M.I.T. Computation Center for providing IBM 704 computer time and other assistance in the initial phase of this study.

Considerable support in the form of card punching, machine programming and IBM 650 computer time were provided by the computer branch of the Geophysics Research Directorate at L. G. Hanscom Field. I appreciate the efforts of Mrs. Janet Dwyer, Miss Dottie Quinn, Mr. Dan La Chance, Mr. Milton Jones, and Airman Carl Collins.

Finally, thanks are due Dr. Norman Phillips, Dr. Barry Saltzman, Dr. Ralph Shapiro, Major David Eddleman, Lt. David Roberts, and Mr. Philip Frank for their cooperation and advice during various phases of this study.

REFERENCES

- Austin, J. M., and L. Krawitz, 50-millibar patterns and their relationship to tropospheric changes, *J. Meteorol.*, **13**, 152-159, 1956.
- Brunt, D., *Physical and Dynamical Meteorology*, Cambridge Univ. Press, London, 1941.
- Charney, J. G., On the scale of atmospheric motions, *Geofys. Publikasjoner, Norske Videnskaps-Akad. Oslo*, **17** (2), 1948.
- Davidson, B., and H. H. Lettau, *Exploring the Atmosphere's First Mile*, vol. I, Pergamon Press, New York, 376 pp., 1957.
- Ekman, V. W., *Nyt, Mag. Naturv.*, **40** (1), 1902.
- Ferrel, W., Essay on the winds and ocean currents, *Nashville J. of Medicine and Surgery*, **11**, 287-301, 375-389, 1856.
- Hansrote, L. S., and J. K. Lambert, Energy interaction between the troposphere and stratosphere, Master's thesis, Mass. Inst. Technol., 1960.
- Kochanski, A., Thermal structure and vertical motion in the lower stratosphere, *Air Weather Serv. Tech. Rept. 105-129*, USAF, Washington, 36 pp., 1954.
- Margules, M., Über die Energie der Stürme, *Jahrb. (Amh. 1903) Zentralanst. f. Meteorol. Geodyn.*, Wien. (English translation in *The Mechanics of*

- the Earth's Atmosphere, a Collection of Translations by Cleveland Abbe, Smithsonian Inst. Misc. Collections*, 51(4), 533-595, 1910), 1903.
- Panofsky, H. A., The effect of vertical motion on local temperature and pressure tendencies, *Bull. Am. Meteorol. Soc.*, 25, 271-275, 1944.
- Phillips, N. A., Energy transformations and meridional circulations associated with simple baroclinic waves in a two-level, quasigeostrophic model, *Tellus*, 3, 273-286, 1954.
- Roberts, D. L., Vertical transport of potential energy across isobaric surfaces, Master's thesis, Mass. Inst. Technol., 1960.
- Rossby, C.-G., The scientific basis of modern meteorology, in *Climate and Man: Yearbook of Agriculture*, U. S. Govt. Printing Office, Washington, D. C., 38-55, 1941.
- Saltzman, B., and A. Fleisher, The modes of release of available potential energy, *Sci. Rept. 1, Contract AF 19(604)-4958*, Mass. Inst. Technol., 20 pp., 1959.
- Starr, V. P., Application of energy principles to the general circulation, in *Compendium of Meteorology*, Am. Meteorol. Soc., Boston, 568-574, 1951.
- Starr, V. P., and R. M. White, A hemispheric study of the atmospheric angular momentum balance, *Quart. J. Roy. Meteorol. Soc.*, 77, 215-226, 1951.
- Starr, V. P., and R. M. White, Schemes for the study of hemispheric exchange processes, *Quart. J. Roy. Meteorol. Soc.*, 78, 407-410, 1952.
- Taylor, G. E., *Phil. Trans. Roy. Soc., London*, 215, 1, 1915.
- White, R. M., and G. F. Nolan, A preliminary study of the potential to kinetic energy conversion process in the stratosphere, *Rept., Dept. of Meteorol., AEC (30-1)2241*, Mass. Inst. Technol., 1959.
- White, R. M., and B. Saltzman, On conversion between potential and kinetic energy in the atmosphere, *Tellus*, 3, 357-363, 1956.
- Wiin-Nielsen, A., A study of energy conversion and meridional circulation for the large-scale motion in the atmosphere, *Monthly Weather Rev.*, 87, 319-332, 1959.

(Manuscript received September 1, 1960;
revised January 10, 1961.)

An Experiment in the Study of the Structure of an Atmospheric System by Numerical Process

STANLEY E. ASPLUND

*Department of Meteorology, Florida State University¹
Tallahassee, Florida*

Abstract. A model nondivergent wind field represents the vorticity-advection structure in a cutoff cold low. Under the hypothesis that the prescribed relative vorticity is steady-state, the scalar vorticity equation and the divergence equation are combined to form an iterative numerical scheme for deduction of the three-dimensional distributions of horizontal divergence, vertical motion, and irrotational part of the wind field. The development and testing of the numerical process are emphasized. On a strictly empirical basis, it is concluded that the computed solution is stable and convergent. This gives a means for mathematical modeling capable of providing explicit numerical information on several matters of interest in connection with equations widely used in the study of atmospheric motion systems.

INTRODUCTION

This report deals with the formulation and solution of a numerical problem concerned with three-dimensional distributions of horizontal divergence, vertical motion, and the irrotational part of the horizontal wind field. This work is in the realm of application of high-speed computing to study of atmospheric dynamics by mathematical modeling. Its purpose is exploration of a possibility, made feasible by fast computing machines, for deduction of space distributions of quantities that cannot be satisfactorily obtained by direct observation.

As a starting point, it is supposed that a given synoptic system may be characterized by its distribution of the vertical component of absolute vorticity. In the case at hand, this premise is expressed by direct prescription of the nondivergent part of the horizontal wind field and the local time derivative of vorticity. The nondivergent part of the wind is an assumed function defined on continuous space coordinates throughout the volume occupied by the system, and the time variation of vorticity is made identically zero by considering a postulated moment in the life cycle of the system when there is a steady state of vorticity.

The problem is to determine the space dis-

tributions of divergence, divergent wind, and vertical motion consistent with the assumed space and time distributions of vorticity, under the constraints of appropriate dynamic equations and boundary conditions. The present specialization of the starting point clearly limits the usefulness of the results for the study of atmospheric systems, but the first objective must be to develop a procedure that works. To do this even for the model system suggested in the preceding paragraph involves extended iterative solution of the nonlinear vorticity equation and the continuity equation. Since this untried numerical method has to be tested empirically for proper behavior, its application to real systems with nonsteady vorticity and imperfectly observed winds belongs in a later stage of investigation.

The present formulation does have meteorological significance, however. The cutoff high-level cyclone is a well-known synoptic model for which the present specializations of the basic information requirements are reasonable. Meaningful results about the structure of cutoff cold lows by the method may be expected. Prescription of an acceptable wind field does not present much of a problem for such a well-defined and symmetric entity [Palmén, 1949]. Furthermore, it is consistent with impressions gained by study of contour-field representations of cutoff cold lows to suppose that they change very slowly. It is, of course, a venture into unexplored territory to assume that the qualitative steadiness of these

¹ Present affiliation: Operations Evaluation Group, Office of the Chief of Naval Operations (Op-03EG), Navy Department, Washington 25, D.C.

systems is correctly expressed as steadiness of the relative vorticity throughout the modeled region at one time. The present work may thus be considered an investigation of the divergence-linked structural characteristics of a model cold low, with uniform steadiness of vorticity as a postulate whose validity is to be indicated by its implications.

It will be seen that the problem at hand has two aspects of experimentation. The present report deals primarily with the first aspect, namely, development of the numerical process. Establishment of the nondivergent part of a model wind field appropriate to a cutoff high-level cyclone and of boundary conditions is essential background. An iterative procedure for deduction of the divergence-linked structural elements is explained and applied. Results of the experiment are interpreted with regard to the behavior of the numerical process. It will be seen that there is every indication of stability and satisfactory coalescence of the solution. It is thus evident that the computed results give an

internally consistent motion-system model complete with divergence and vertical motion. Its horizontal wind field is accurately decomposed into irrotational and nondivergent parts.

The second aspect of experimentation concerns what may be learned about the structure of the particular synoptic entity which provides the numbers used in exploration of the numerical process. Although it is not within the chosen scope of this paper to deal with meteorological implications explicitly concerning the structure of cold lows, some points concerning the computed constitution of the scalar vorticity equation are examined in the context of the model.

THE NONDIVERGENT WIND MODEL

General form. Selection of an idealized prescription of the nondivergent wind in a cutoff low at maturity was based as much as possible on observed cases. It is found that representation of such a system by combination of a circularly symmetric vortex and zonal flow is quite acceptable. In the mid-troposphere, at least

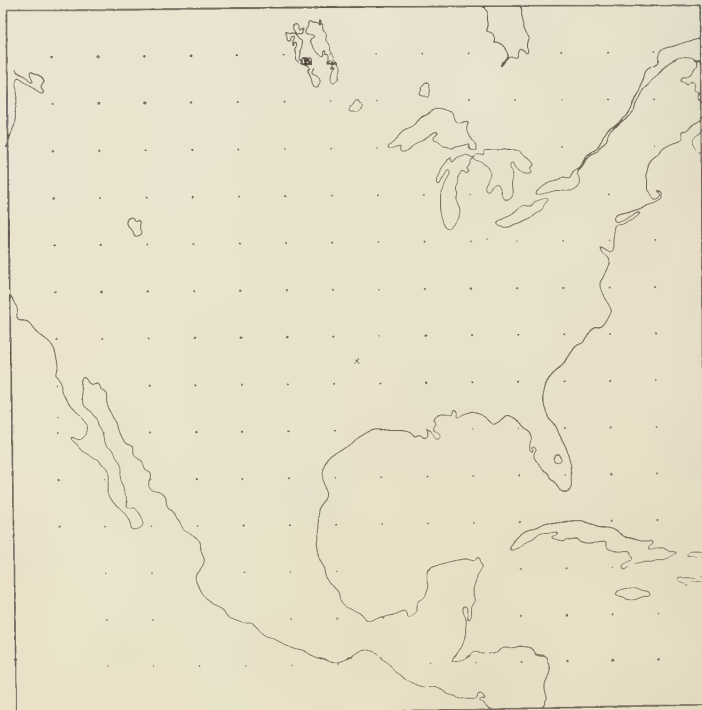


Fig. 1. Region of computation and finite difference grid.



Fig. 2. Upper: Radial profile of vortex components of model wind field on reference surface. Lower: North-south profile of zonal component.

such a model is satisfactory, even considered part from the need for simplicity of the assumed nondivergent field in developing the numerical process.

The model horizontal velocity field has the form

$$\mathbf{V}(x, y, p) = G(p)\mathbf{c}(r) + F(p)\mathbf{U}(Y) \quad (1)$$

The variables (x, y) refer to a Cartesian coordinate system with its origin at latitude 34°N on a Lambert conformal conic projection. For convenience of assignment of the specific functional expressions, the vortex part \mathbf{c} is given in terms of polar coordinates (r, θ) on the map, and the zonal part \mathbf{U} is a function of the distance Y measured from 34°N along the map meridian.

Assigned wind-field structure on pressure surfaces. The vortex wind \mathbf{c} has no radial component, and depends only upon the radial coordinate. The distance dimension for the magnitude c of the vortex speed is meters on the meteorological coordinate curves, but the units of r are map centimeters. Inspection of the polar-coordinate expression for the divergence of a horizontal vector shows that $G(p)\mathbf{c}(r)$ is nondivergent. The divergence of the zonal part of the basic model is also readily seen to be zero, by inspection of its expression in the meteorological coordinate system [Haltiner and Martin, 1957]. As in the case of the vortex part of the wind field, conversion of the map distance units of Y

to earth distance units has no effect on the conclusion that the model is nondivergent.

The frame of reference for subsequent discussions is given by Figure 1. This figure shows the part of North America and its surroundings over which the wind field is prescribed. A finite-difference grid with 14 rows and 14 columns of interior points is indicated by dots, and the origin of the (x, y) coordinate system is marked by a circled cross. It is convenient at times to refer to the coordinates r and Y in terms of centimeters on the map. For orientation in such references, one may note that the spacing of grid points is 4 cm. This spacing corresponds to 300 km at the standard parallels 25°N and $48^\circ30'\text{N}$. The format of interior grid point identification is $ijjj$, $01 \leq ii \leq 14$ being the row number and $01 \leq jj \leq 14$ the column. It may be helpful to observe that 1° of latitude is approximately 1.5 cm on the map used.

The vortex speed $c(r)$ is modeled by adjustment of the parameters a_1 , a_2 , and a_3 in the sum of two functions of the form $\exp[-a_1(z - a_2)^2/a_3]$. In this case the dummy variable z is the map coordinate r . The zonal wind speed $U(Y)$ is a single function of the same form, with z meaning Y . Numerical values are set forth in an earlier publication [Asplund, 1959]. The vortex and zonal wind speed profiles are shown in Figure 2. The shapes of these curves are applicable for the respective parts of the assumed total wind on

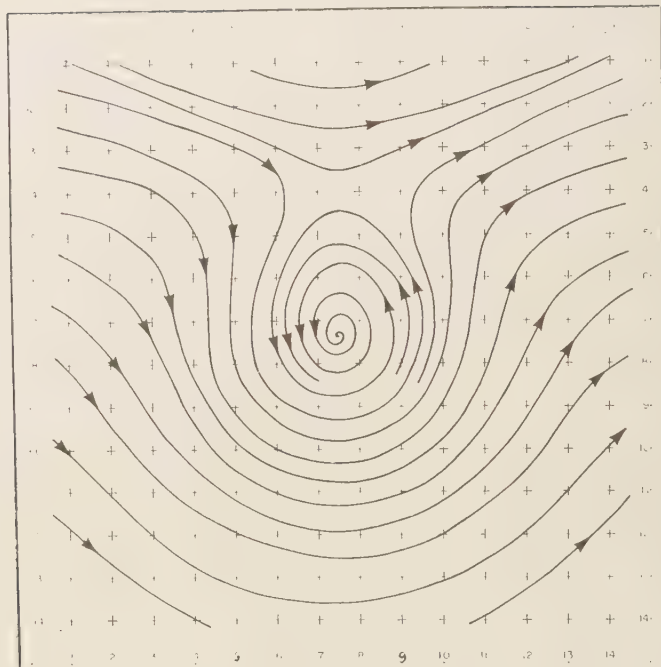


Fig. 3. Streamline analysis for V_T at 500 mb.

any pressure surface. The scale shown applies in the mid-troposphere, at a reference pressure surface for which the functions $F(p)$ and $G(p)$ in equation 1 have the value of unity. It was at first intended that the reference surface would be at 500 mb. In the course of adjustments to get better resemblance to certain observed prototype cases it became convenient to put this reference level at 565 mb.

The vortex wind profile for the reference level has a maximum speed nearly 30 m sec^{-1} at a radial distance of $6\frac{1}{4}$ latitude degrees from the center. The profile approaches zero asymptotically for large r . The zonal profile for the reference level reaches a maximum of 26 m sec^{-1} at 52°N , which is 18 latitude degrees north of the vortex center. The zonal speed drops off to the south of the maximum in a flat Gaussian error curve. The zonal speed at the latitude of the vortex center is just over 5 m sec^{-1} at the reference level.

Streamline representation of the nondivergent wind field at 30 cb as given by combination of c and U may be seen in Figure 3. The data for

this analysis include the computed divergent part of the wind field, but the differences from the nondivergent field are unimportant for the present purpose. Figure 4 is the analysis of the speed field for V_T at 50 cb. The axis of maximum wind around the vortex comes out as a very stylized representation of the characteristic localized and traveling 'jet streams' often found in systems of this kind.

Vertical variation of nondivergent wind. The vertical structure of the model wind field is determined by assignment of values to the functions $F(p)$ and $G(p)$ at 100-mb intervals shown in Table 1. These values were obtained by experimentation and comparison of the model to observed cases, in an effort to make the vertical structure realistic. Their specification and the assumption of a structure representable by a circular vortex in a zonal current determine the dynamic consistency of the model. It will be appropriate later on to examine the authenticity of the distributions of mass and energy built into the nondivergent model by these assumptions.

Graphs showing the vertical profiles of the

maximum wind in the two components of the nondivergent field, plotted against $\ln p$, are shown in Figure 5. The magnitudes of the maximum vertical shears at 500 mb are 4.6 m sec^{-1} per kilometer in the vortex, and 5.0 m sec^{-1} per kilometer in the zonal current. Graphs similar to Figure 5 were used in obtaining the derivatives of the vertical variation functions by measurements of slope. Observed cold-low cases were treated in terms of such graphs as a control on the assignments of $F(p)$ and $G(p)$ for the model.

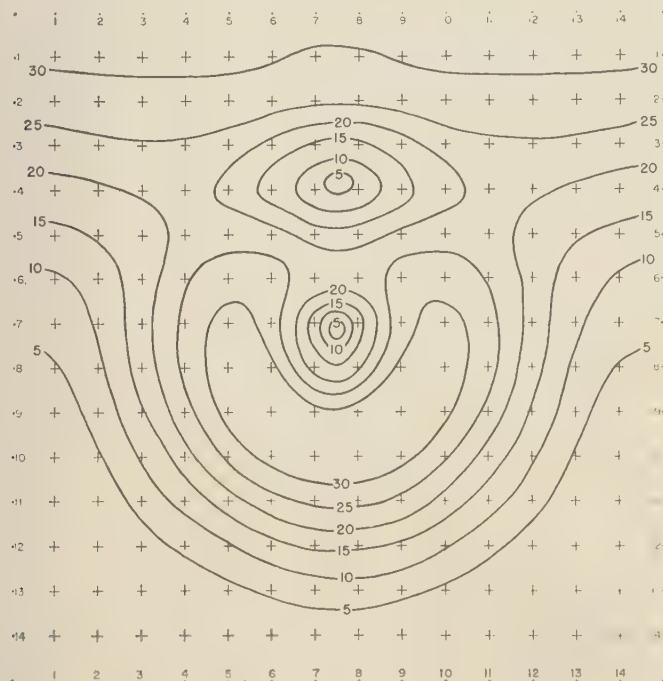
Vorticity of model. Underlying the methodology under development here is the thought that it is feasible to characterize atmospheric motion patterns in terms of vorticity fields. The difficulties which will attend putting such a scheme into effect are many. The application at hand avoids these problems by assuming a nondivergent wind field with steady vorticity as the subject for investigation. A degree of allegiance to a vorticity-pattern basis of classification is maintained by giving the vorticity patterns here as part of the basic presentation of the model. The upper part of Figure 6 shows the radial

TABLE 1. Values of Vertical Variation Functions and First Derivatives

p , centibars	$F(p)$	$G(p)$	$F'(p)$, cb^{-1}	$G'(p)$, cb^{-1}
90	0.320	0.500	-0.0158	-0.0120
80	0.490	0.625	-0.0178	-0.0128
70	0.680	0.760	-0.0204	-0.0153
60	0.910	0.930	-0.0250	-0.0180
50	1.180	1.140	-0.0280	-0.0210
40	1.470	1.360	-0.0300	-0.0230
30	1.770	1.480	-0.0300	-0.0280

distribution of relative vorticity of the reference-surface vortex $c(r)$. The lower graph shows the variation of relative vorticity of the complete nondivergent wind along the north-south axis on the reference surface. These vorticity values were obtained by differentiation of $c(r)$ and $U(Y)$ with provision for conversion of map coordinate system increments in r and Y to distances on the mapped surface. Thus

$$\zeta_c = m \, dc/dr + \bar{m}c/r \quad (2)$$

Fig. 4. Isotach analysis for V_T at 500 mb. Units m sec^{-1} .

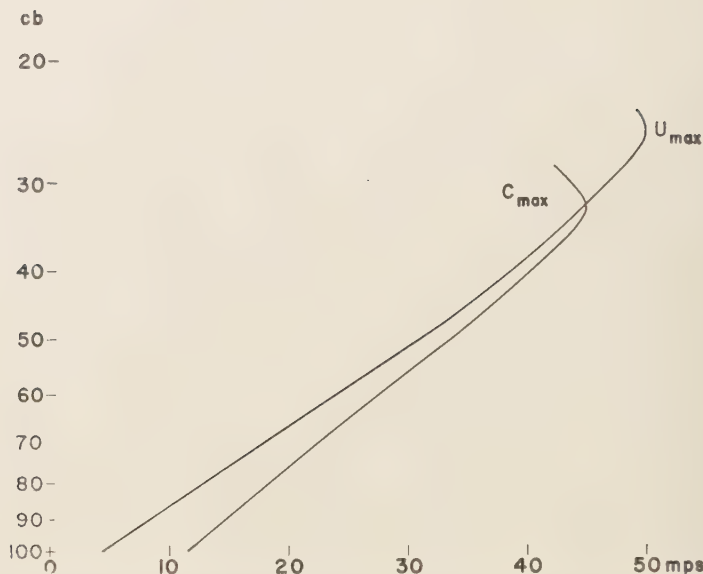


Fig. 5. Vertical profiles of maximum wind speed in vortex (C_{\max}) and zonal current (U_{\max}).

and

$$\zeta_U = -m \, dU/dY + (U \tan \phi)/a \quad (3)$$

gave the vortex relative vorticity and zonal relative vorticity, respectively. Here a is the

mean radius of the earth and m is the map-scale factor defined by

$$ds/ds' = m \quad (4)$$

where ds' is distance in meters on the map.

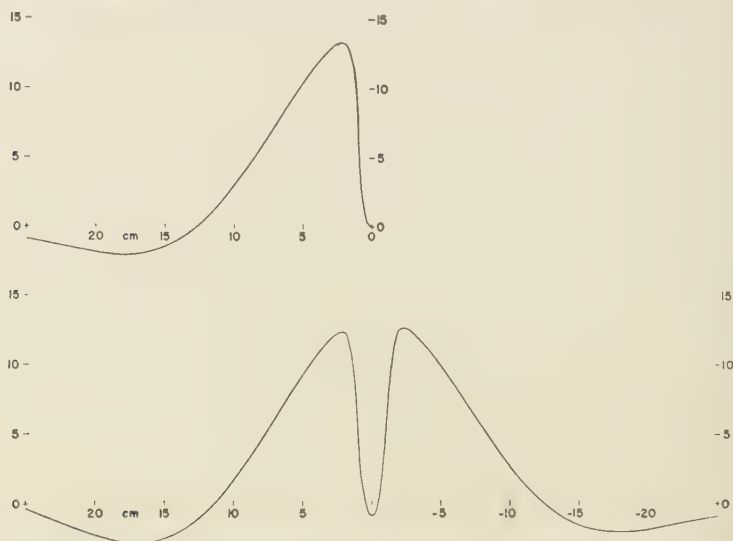


Fig. 6. *Upper:* Radial profile of relative vorticity of vortex at reference level. Units 10^{-5} sec^{-1} . *Lower:* North-south profile of relative vorticity for combined zonal current and vortex. North is at left. Units 10^{-5} sec^{-1} .

surface and ds is centimeters on the map. The symbol \bar{m} in (2) stands for a mean value approximated to small error by the arithmetic average of the values at 34°N and the latitude of the point of evaluation.

It is to be noted that prescription of the nondivergent wind model by tabulated functions defined on the continuous space coordinates obviates finite-differencing errors in computing vorticity and its advection by the nondivergent wind.

The vorticity of the zonal wind field is very small. The extreme value is $-1.3 \times 10^{-6} \text{ sec}^{-1}$ at 41°N . It is anticyclonic south of 50°N and cyclonic elsewhere. At the latitude of the vortex center it is only $-0.9 \times 10^{-6} \text{ sec}^{-1}$. The vorticity field of the total model is overwhelmingly determined by the vortex component. This has an annular ring of maximum at 150 km from the center, where the curvature term is very important. The shear and curvature terms are both positive out to the radial distance of the peak vortex wind, about 700 km. Farther out the shear is anticyclonic, accounting for anticyclonic rela-

tive vorticity at large radial distances, e.g., 950 km at the latitude of the vortex center.

The distribution of absolute vorticity at 500 mb is shown by a plan-view analysis in Figure 7. The maximum of absolute vorticity is 22×10^{-6} in a roughly annular pattern of 150-km radius. Similar patterns of absolute vorticity with smaller extremes obtain at lower levels. The magnitude in the annular maximum for the 900-mb surface is $14 \times 10^{-6} \text{ sec}^{-1}$.

The pattern of relative vorticity is easily visualized by consideration of Figure 7 and the values of the Coriolis parameter, and also from the profiles in Figure 6. It has a zero value very near the center of the vortex and the annular maximum as in the absolute vorticity pattern.

OUTLINE OF METHOD

A nondivergent wind field with steady-state vorticity is now known through a certain volume of troposphere. The boundaries in the partial-differential-equation system for further deduction of structure are a certain lower constant-pressure surface and four vertical planes.

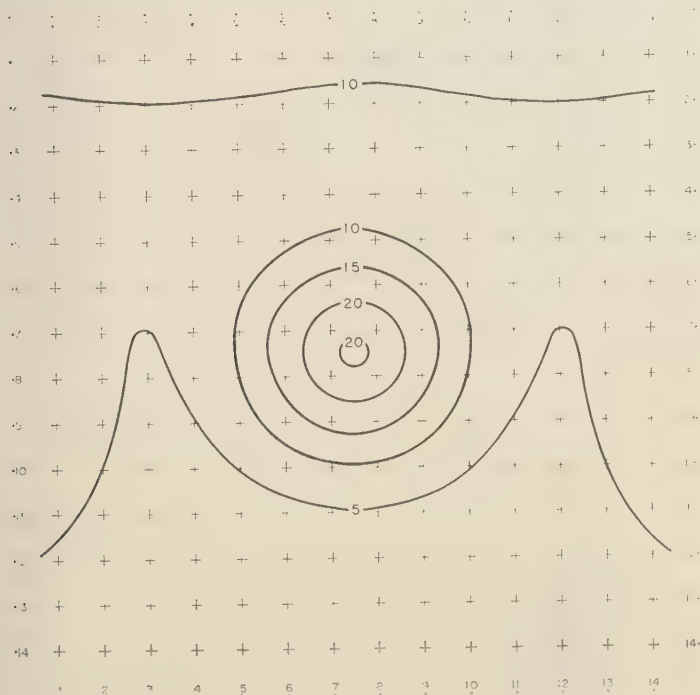


Fig. 7. Distribution of absolute vorticity on 500-mb surface. Units 10^{-6} sec^{-1} .

The problem is to obtain the total wind field and the substantial derivative of pressure in the region. The total wind field of the system, denoted by \mathbf{V}_T , is the sum of the nondivergent wind \mathbf{V} and an irrotational part \mathbf{W} , called the 'divergent wind.' Thus

$$\mathbf{V}_T = \mathbf{V} + \mathbf{W} \quad (5)$$

where $\nabla \cdot \mathbf{V} = 0$. We note also that

$$\mathbf{k} \cdot \nabla \times \mathbf{V} = \zeta = \zeta_T \quad (6)$$

is the vorticity of the model nondivergent wind and the total wind field as well. This is so because

$$\nabla \times \mathbf{W} = 0 \quad (7)$$

by definition. The separation (5) is an application of a well-known theorem of vector analysis [Phillips, 1933] to the two-dimensional wind field on each pressure surface.

The deduction of \mathbf{W} to complete our knowledge of the total wind field is made possible by the irrotational property of \mathbf{W} and the nondivergent property of \mathbf{V} . By the latter we have

$$\nabla \cdot \mathbf{V}_T = \nabla \cdot \mathbf{W} = D \quad (8)$$

The symbol D for the divergence of \mathbf{W} is hereby introduced. Solution for \mathbf{W} is based on making (7) and (8) into a pair of Poisson equations for the components of \mathbf{W} [Sherman, 1952].

To do this, the constraints exercised by the governing equations of atmospheric dynamics come into the picture. The key to the whole procedure is use of the scalar vorticity equation for determination of the divergence D . Under the assumptions of steady-state vorticity and the separation (5) with its attendant properties (6) and (8), the scalar vorticity equation in the p -coordinate system [Eliassen, 1949] becomes

$$D = -(\zeta + f)^{-1} \cdot \left[\mathbf{V} \cdot \nabla (\zeta + f) + \mathbf{W} \cdot \nabla (\zeta + f) + \mathbf{k} \cdot \nabla \omega \times \frac{\partial \mathbf{V}}{\partial p} + \omega \frac{\partial \zeta}{\partial p} \right] \quad (9)$$

The conventional symbol ω introduced here for the substantial derivative of pressure will henceforth be called 'vertical motion,' when convenient.

The above form of the vorticity equation is nearly exact. The only approximations are neglect of viscosity and a term expressing the

effect of differential vertical advection of \mathbf{W} momentum [Sherman, 1952]. With these omissions, no nonlinear terms are involved in the present use of (9).

The system of equations is completed by inclusion of the continuity equation

$$\partial \omega / \partial p = -D \quad (10)$$

It may be observed that the use of the vorticity equation to obtain the divergence, somewhat as in (9), followed by integration of (10) is a version of the 'vorticity method' for determination of vertical motion. An important difference is that in the usual vorticity method the advection of absolute vorticity by the divergent wind is not treated separately and all vertical motion effects in (9) are neglected. Sometimes vertical advection is treated but not the tipping term [Deardorff, 1957].

It is now proposed to solve equations 7 through 10 for D , \mathbf{W} , and ω by an iterative process of successive approximations. The process is begun by taking as a first approximation to D the value that balances the first term on the right side of (9):

$$D^{(0)} = -(\zeta + f)^{-1} \mathbf{V} \cdot \nabla (\zeta + f) \quad (11)$$

This term is completely known by assumption of the nondivergent wind field. It is the major part of the divergence necessary to satisfy the vorticity equation. In application of the usual vorticity method to a system having small local rate of change of vorticity, (11) would be the whole answer.

Most of the computation involved in getting $D^{(0)}$ is the same for all pressure surfaces. It was first carried out for the reference surface, where $F(p)$ and $G(p)$ are unity. The expansion of the right side of (11) in terms of the winds $c(r)$ and $U(Y)$, the vorticities ζ_c and ζ_U , their gradients and the map-scale factor and other parameters may be found in an earlier publication [Asplund 1959].

A plan of iteration having the following framework is proposed as a way of arriving at values of divergent wind, vertical motion, and divergence satisfying (9):

$$D = D^{(0)} + \Delta^{(H)} D + \Delta^{(V)} D \quad (12)$$

The initial approximation $D^{(0)}$ is the basis for computation of a horizontal divergence-linked

crement of divergence $\Delta^{(H)}D$. The improved value of divergence obtained by applying the H -type correction then goes into the computation of a vertical divergence-linked increment $\Delta^{(V)}D$. Alternating application of H -type and V -type increments produces a sequence of corrections of each type, and a corresponding sequence of approximations to D . There are a number of possible variations on this general plan. The order of application of V - and H -type corrections may be reversed, and the two types of corrections need not be applied in one-and-one alternation. To make explicit the connection between (12) and (9), the H -type increment is produced by correction of absolute vorticity by the divergent increment:

$$\Delta^{(H)} D = -\zeta_a^{-1} \mathbf{W} \cdot \nabla \zeta_a \quad (13)$$

The evaluation from the approximate divergence on one hand is primarily the problem of solving the two Poisson equations that arise from cross differentiation of equations 7 and 8. These equations, written in terms of the components u and v of \mathbf{W} in the rectangular (x, y) grid system, are

$$\nabla^2 u = \frac{\partial D}{\partial x} \quad \nabla^2 v = \frac{\partial D}{\partial y} \quad (14)$$

Each of these is solved subject to boundary conditions discussed in the next section. The method of solution is the extrapolated Liebmann relaxation process. Finite-difference approximations are used on the latest available approximation of D to provide the input distributions of the Laplacians of u and v . The finite-difference approximations to (14) make use of the Cartesian properties of the grid, and so the length units of the solutions u and v are those of the image surface.

The components of u and v in the Cartesian grid system having been obtained, evaluation of (3) in units consistent with those of $D^{(0)}$ is a minor modification of the computation scheme for $D^{(0)}$ [Asplund, 1959].²

The V -type increment comprises the vertical-motion terms in (9), as follows:

² The expressions for $\Delta^{(H)}D$ and $\Delta^{(V)}D$ given as equations 42 and 45 in Asplund [1959] have errors of omission of the map distortion factor, which would divide the term $2\Omega a^{-1} \cos \phi$ in (42) and multiply the right side of (45).

$$\Delta^{(V)} D = -\zeta_a^{-1} \mathbf{k} \cdot \nabla \omega \times \frac{\partial \mathbf{V}}{\partial p} - \zeta_a^{-1} \omega \frac{\partial \zeta}{\partial p} \quad (15)$$

The evaluation of this contribution to a new D at any stage of approximation to D is inexact, because ω is approximate except at the lower boundary. The source of the boundary values $\omega(p_0)$ is discussed in the next section. At subsequent pressure steps ω is obtained by numerical integration of the continuity equation over steps $\Delta p = 10$ cb. It is possible to use centered-difference integration from the outset, by taking advantage of the dominance of $D^{(0)}$ in determination of a value of D at the end of the forward step from p to $p + \Delta p$. The step from 90 cb to 80 cb was made by Euler's modified method [Kunz, 1957]. Subsequent pressure steps were made by Simpson's numerical integration formula (*ibid.*), applied as follows:

$$\omega(p_i) = \omega(p_{i-2})$$

$$+ [D(p_{i-2}) + 4D(p_{i-1}) + D(p_i)] \Delta p / 3 \quad (16)$$

The subscripted pressure values specify the arguments in the sequence 90, 80, 70, 60, 50 cb. The horizontal grid location is the same for all quantities in (16). The values of D and ω for preceding pressure steps are those of the latest stages available.

The summaries in Table 2 clarify the symbolism used in presentation of results and give further explanation of the scheme. Table 2a shows the scheme found by experimentation to give comparable-sized terminal increments of the two types V and H . The procedure is arranged in terms of cycles and steps within cycles, to facilitate discussion. Table 2b indicates variations on the preferred scheme of Table 2a, as used in the experiment.

TABLE 2a. Summary of Numerical Process

Initial Divergence Approximation	Quantities Evaluated	Type of Increment	Cycle and Step
$D^{(0)}$	$\mathbf{W}^{(1)}, \Delta^{(1)}D$	H	I-1
$D^{(1)}$	$\omega^{(1)}, \Delta^{(2)}D$	V	II-1
$D^{(2)}$	$\mathbf{W}^{(2)}, \Delta^{(3)}D$	H	II-2
$D^{(3)}$	$\mathbf{W}^{(3)}, \Delta^{(4)}D$	H	II-3
$D^{(4)}$	$\omega^{(2)}, \Delta^{(5)}D$	V	III-1
$D^{(5)}$	$\mathbf{W}^{(4)}, \Delta^{(6)}D$	H	III-2
$D^{(6)}$	$\mathbf{W}^{(5)}, \Delta^{(7)}D$	H	III-3
$D^{(7)}$	$\mathbf{W}^{(6)}, \Delta^{(8)}D$	V	IV-1

TABLE 2b. Iterations at Individual Pressure Surfaces

Pressure, cb	Steps Completed	Final Increment	Final D
90	V, H, H	$\Delta^{(4)}D$	$D^{(4)}$
80	V, V, H, H	$\Delta^{(4)}D$	$D^{(4)}$
70	I, II, III-1	$\Delta^{(6)}D$	$D^{(6)}$
60	I, I, III	$\Delta^{(7)}D$	$D^{(7)}$
50	I, II, III, IV-1	$\Delta^{(8)}D$	$D^{(8)}$
40	I, II, III	$\Delta^{(7)}D$	$D^{(7)}$
30	I, II	$\Delta^{(4)}D$	$D^{(4)}$

The procedures followed at the 90- and 80-cb surfaces do not fit well into the final scheme indicated by Table 2a. The procedure for the lower boundary surface is necessarily different because the V -type increment is given by the wind prescription and boundary values. The numbering of increments for 90 cb is not consecutive, one number having been skipped to carry out the final system. The second pressure step shows that the idea of having two V -type increments in a row was tried. It was found that the second increment was not significant.

Iteration at the third pressure surface, 70 cb, used the scheme of Table 2a up to the first step of cycle III. Data presented in a later section show the behavior of increments and the need for increasingly greater numbers of steps as the basic nondivergent wind increases, that is, at pressure surfaces higher up in the modeled region.

BOUNDARY CONDITIONS

Vertical motion at the lower boundary. It is assumed that the integrated effect of the divergence of an Ekman-spiral friction layer [Brunt, 1939] gives an adequate lower-boundary value of vertical motion for the model.

The vertical motion at the top of the friction layer is found to be proportional to the geostrophic vorticity [Charney and Eliassen, 1949]. Reversing a common assumption, the relative vorticity computed from the wind is used instead of the geostrophic vorticity. Thus the following formula for the 900-mb distribution of ω arises:

$$\omega = -\rho g(K/2f)^{1/2} \sin 2\theta \zeta \quad (17)$$

Here ρ is a mean value of density for the layer, f is the Coriolis parameter, which has been held constant, and K is the eddy viscosity. The angle θ between the wind and the contour lines at

anemometer level is also considered constant in the horizontal. A further simplification inherent in this device to obtain a lower-boundary prescription is to neglect any divergence of the component of wind along the contours in the friction layer. The assumptions of the Ekman spiral derivation are in addition to those enumerated above.

A numerical value by which to multiply the relative vorticity to obtain the desired grid distribution of ω for 90 cb was obtained by assigning the values $\rho = 1.22 \times 10^{-8}$ mts, $K = 5.5$ mts, $f = 9 \times 10^{-5} \text{ sec}^{-1}$, $\theta = 0.5$ radian. Figure 8 is a scalar analysis of the pattern.

Nondiverging wind at lateral boundaries. Appropriate selection of boundary conditions for \mathbf{W} is important in this experimentation to illuminate the composition of the vorticity-equation terms involving divergent wind and vertical motion relative to those commonly considered of greater importance.

It is pertinent to review briefly certain constraints on freedom of choice of boundary value for the irrotational wind. The results are based on properties of harmonic functions [Phillips, 1933].

The problem of finding \mathbf{W} is essentially that of putting into use the vector-analysis theorem that states (with unimportant restrictions added to the hypothesis for simplicity) that a vector field that is continuous and differentiable in a finite region is completely determined by its divergence, vorticity, and the normal component on the boundary. By the duality of divergence and vorticity of a two-dimensional vector, the tangential component on the boundary may be used instead of the normal. This theorem gives a sort of counterpart to the Helmholtz resolution expressed by equation 5. In going from given divergence and vorticity distributions to the total horizontal wind as a sum of irrotational and solenoidal parts, the boundary condition determines how much irrotational nondivergent wind there is in the result, and how it is partitioned between the components of the Helmholtz resolution.

In the general two-dimensional problem in which the vertical component of vorticity and the divergence are given and the total wind is to be determined, one commonly solves Poisson equations for the potential function that exists for the irrotational part and the stream function

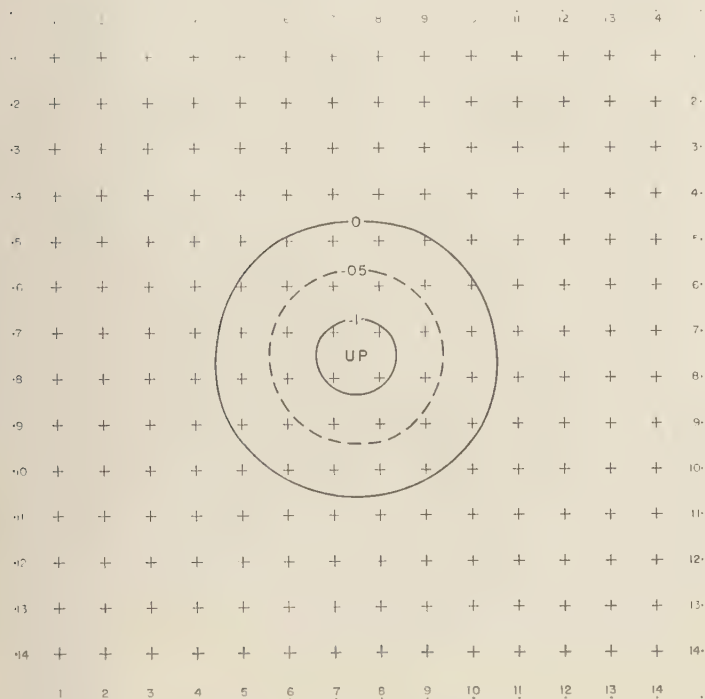


Fig. 8. Distribution of ω on 90-cb surface. Units 10^{-4} cb sec $^{-1}$.

that exists for the nondivergent part. The Laplacians of these functions are equal to the given divergence and vorticity, respectively, in this standard approach. The respective parts of the total wind field are then obtained by differentiating these scalar functions. Alternatively, Poisson equations that equate the Laplacians of the unknown velocity components to combinations of derivatives of the known divergence and vorticity are solved; this is the procedure followed in solving for \mathbf{W} in this problem. The condition of zero vorticity merely specializes the problem, including the considerations relevant to the effect of boundary conditions. It is convenient to talk about these effects in terms of the potential function and stream function. The reason for this is essentially that, when potential and stream functions exist together for a part of the wind field, they are harmonic functions.

The various alternatives for boundary conditions on the potential and stream functions lie between two extremes. On the one hand, the stream function may be required to be constant

on the boundary. This is equivalent to saying that the nondivergent part of the wind field has no normal component. It is convenient to do this when the given boundary condition is the normal component of the total wind [Eliassen, 1955]. If we add the assumption that the boundary value of the part of the stream function that gives rotational nondivergent wind is also a constant on the boundary, it follows that the rest of this stream function is also a constant on the boundary. But the remaining part of the stream function gives irrotational nondivergent wind, hence is a harmonic function (necessary continuity being assumed), hence is constant throughout the region. This means that none of the nondivergent wind is irrotational. The same assumptions give the result that the area integral of kinetic energy for the nondivergent wind field is the sum of the kinetic energies of its rotational and irrotational parts [Petterssen and Calabrese, 1959]. Hence the interpretation may be made that setting the stream function constant on the boundary maximizes the kinetic energy of the

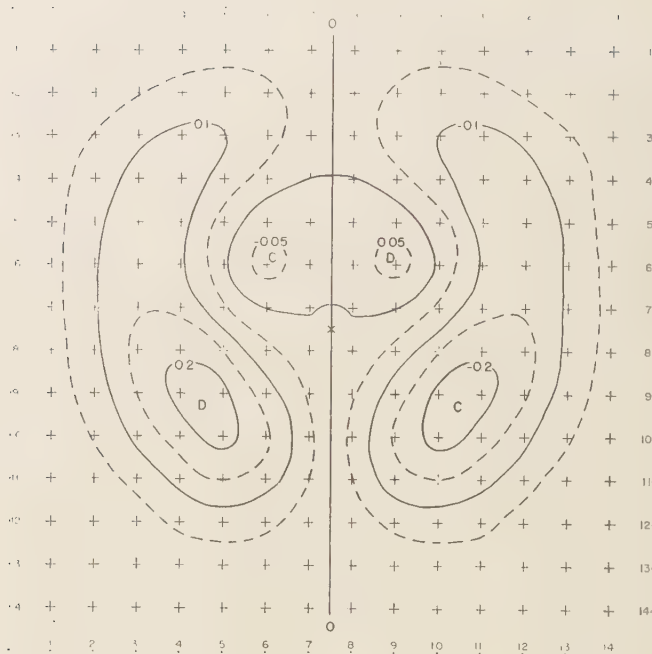


Fig. 9. Divergence approximation $D^{(0)}$ on 90-cb surface, units 10^{-5} sec^{-1} .

part of the wind field that has a given divergence.³

The second possibility for a sufficient boundary condition is knowledge of the tangential component of the total wind. In this case it is convenient to assume that the potential function is constant on the boundary, thereby making the tangential component of the corresponding part of the total wind field be zero. Considerations similar to those discussed above show that this has the effect that any irrotational nondivergent wind called for by the boundary condition and interior Laplacian value is put into the part of the wind having the given vorticity.

In the problem at hand, the application of these ideas is made in terms of the two-dimensional wind vector \mathbf{W} . As shown by equations 5, 7, and 8, this wind is to have zero vorticity over the horizontal grid, and divergence D

successively approximated from equation 9. Equations 14 express these requirements in Poisson equations for the components u and v of \mathbf{W} .

It is intended that any nondivergent irrotational part of \mathbf{V}_T shall be in the model wind \mathbf{V} . That is, there is to be nothing in the solution of the equations 14 other than the irrotational wind having the divergence D .

The solution may still be thought of as consisting of a part having a potential, hence is irrotational, and a nondiverging part having a stream function. Application of boundary considerations directs that the tangential boundary component of the solution be zero, to eliminate nondivergent wind from the part of the solution having a potential.

In the absence of any reason for doing otherwise, the conditions set on equations 14 were that, at the boundary B ,

$$\mathbf{W}(B) = 0 \quad (18)$$

The implication that the normal boundary component is zero means that the solution of (14) has no nondivergent irrotational part at all.

³ It is proved in the paper by *Petterssen and Calabrese* [1959] that the hypotheses (a) that kinetic energy of the nondivergent wind is minimum and (b) that values of the stream function for the nondivergent part having the given vorticity are constant on the boundary imply constancy of the rest of the stream function in the whole region.

It is to be noted that imposition of both tangential and normal boundary component is overspecification. By specifying the normal boundary component, the part of the solution given by the potential function is not left free to satisfy the integral condition of Gauss's divergence theorem. That is, if only a tangential boundary condition is imposed, the normal component $\mathbf{W} \cdot \mathbf{n}$ yielded by solution of (14) would come out such that

$$\int_B \mathbf{W} \cdot \mathbf{n} \, ds = \iint_R D \, d\sigma \quad (19)$$

where ds is an element of boundary distance and $d\sigma$ is an element of area of the region of integration R .

It is believed that the overspecification involved in using the boundary condition (18) does not produce significant error. The antisymmetry of the divergence pattern is perfect for $D^{(0)}$ and not discernibly unbalanced for higher approximations to D . Consequently, the area integral of divergence is so nearly zero that the integral condition (19) could hardly be better satisfied

by more correct (and cumbersome) determination of the normal boundary component of \mathbf{W} .

RESULTS OF THE EXPERIMENT

Vorticity-advection structure of the nondivergent wind model. The zero-order estimate of the divergence required to maintain steady-state vorticity in the assumed wind model is shown for 90, 70, and 50 cb in Figures 9, 10, and 11. There is a center of divergence in the near northeast quadrant and a lobe of convergence farther out. Corresponding patterns with opposite signs occur in symmetrical locations. The essential similarity of the model wind streamlines and vorticity isopleths at the different pressure surfaces is evident in the similarities of $D^{(0)}$ patterns shown in these figures. There is, however, a discernible change in the composition of the net vorticity advection pattern as the zonal and vortex wind fields and their associated vorticities vary with height. The magnitude of $D^{(0)}$ at 90 cb is $2 \times 10^{-6} \text{ sec}^{-1}$ in a region where the ' βv ' term is prominent. On the 50-cb surface a magnitude 5 times as great can be found, more than 10^{-5} sec^{-1} . Such values occur in extrema

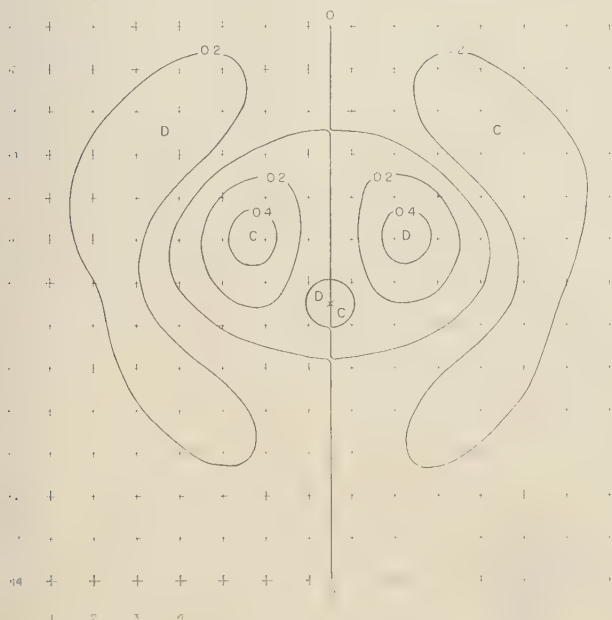


Fig. 10. Divergence approximation $D^{(0)}$ on 70-cb surface, units 10^{-5} sec^{-1} .

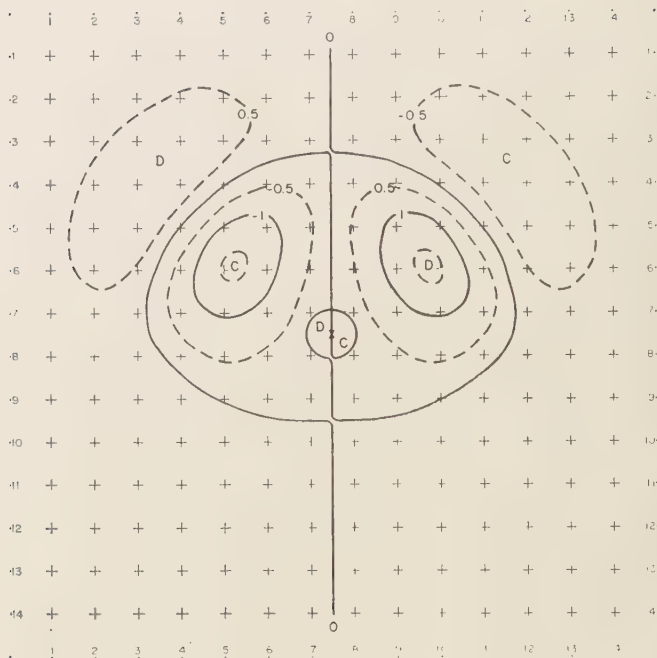


Fig. 11. Divergence approximation $D^{(0)}$ on 50-cb surface, units 10^{-5} sec^{-1} .

located nearer the center of the system, in which the effect of advection of vortex vorticity by the zonal wind increases with height at a greater rate than advection of planetary vorticity (βv).

It is important to understand the anatomy of the $D^{(0)}$ pattern. The patterns of divergence distribution for all pressure surfaces in the modeled region are evident from the examples of $D^{(0)}$ given in Figures 9, 10, and 11. It is clear that assuming \mathbf{V} to have steady-state vorticity implies that compensation for tropospheric pressure change would have to occur in the stratosphere, with the sign change of divergence occurring very much higher than in the typical trough-ridge system [Fleagle, 1947; Arnason, 1955]. There are suggestions in the literature that development of troposphere-deep layers of one sign of divergence takes place as systems reach maximum intensity [Petterssen, 1956]. In this model we have it to a startling degree. The correspondence between the model and reality has not been investigated well enough for discussion here. It may be noted, however, that prescription of the nondivergent wind was based on careful study of observed cutoff cold lows.

Qualitative reconsideration was given to the model wind field, after it was seen that the divergence pattern was not much changed from $D^{(0)}$ by the divergent-wind advection and vertical motion terms. This reconsideration still showed no reason to contemplate any materially different basic structure for a cutoff cyclone model. Further information on this point was anticipated from a sequel investigation using the divergence equation to compute the geopotential field of the system. Any lack of verisimilitude of the model to reality is probably due to the assumption of zero vorticity tendency everywhere at once.

Table 3 shows the constitution of the quantity $D^{(0)}$ on the reference surface. Values are given for enough of the northeastern quadrant to bring out the main point. The nature of the model is such that there are three contributions to $D^{(0)}$. There is advection of the vorticity of the zonal flow by the vortex wind field ($\mathbf{c} \cdot \nabla \zeta_u$), advection of the vorticity of the vortex wind field by the zonal flow ($\mathbf{U} \cdot \nabla \zeta_v$), and the advection of planetary vorticity ($\mathbf{c} \cdot \nabla f$).

Examination of the tabulated quantities

reveals that the principal extrema of the $D^{(0)}$ field (see, for example, grid point 0610) are due to advection of the cyclonic vorticity of the vortex by the zonal wind field. The key point of the geometry that brings this about is that the gradient of vortex vorticity ($\nabla \zeta_c$) is directed radially inward for any point between the ζ_c maximum at 150 km from the vortex center and the minimum about 1250 km out. The only way that the divergence center in the vicinity of point 0610 on all pressure surfaces of the region could be substantially altered is by greatly changing the zonal wind. For example, if zonal wind were from the east relative to the vortex, the signs of $D^{(0)}$ would be reversed. The nature of the $\mathbf{U} \cdot \nabla \zeta_c$ contribution is an important consideration because of its importance in determining $D^{(0)}$. Further, we see by the agreement of sign of these quantities that, over most of the region in which the zonal-wind advection of vortex vorticity is the main effect, the advection of zonal-wind vorticity by the wind of the vortex augments it.

Coalescence of divergence values at 50 cb. The successive approximations to the divergence for the 50-cb surface are summarized and discussed in this section from the point of view of their empirical indication of truncation error. The numerical scheme replaces the divergence required by the vorticity equation with the value attained after a few cycles, as shown in Table 2. Sample data from the 50-cb surface have been chosen to show the behavior of the series at any of the constant-pressure surfaces. One selection of values is given in Table 4. The grid points of this tabulation include those for which there occurred an extreme value of the initial divergence $D^{(0)}$, the final H -type or final V -type increment, or the last change of either type of increment. The purpose of the table is to indicate the sizes of the successive steps toward the finally accepted divergence. As a by-product, several properties of the process are indicated by these sample values and their average absolute values in the last line of Table 4. It is apparent that the first of each of the H and V types of correction to $D^{(0)}$ is not much changed by successive refinements. To eliminate reference to Table 2 for review of the notation, we note that the divergence increments with superscripts 1, 3, 4, 6, and 7 are horizontal-type, and those with superscripts 2, 5 (and 8) are vertical-type. The data for $\Delta^{(8)}D$ and $D^{(8)}$ are not included in the

TABLE 3. Vorticity Advection Contributions of Zonal and Vortex Winds

(Grid Point	$D^{(0)}$, 10^{-5} sec $^{-1}$	$\mathbf{c} \cdot \nabla \zeta_c$		$\mathbf{U} \cdot \nabla \zeta_c$		$\mathbf{c} \cdot \nabla f$	
		sign	%	sign	%	sign	%
0308	-0.08	-	48	-	8	-	44
0309	-0.30	-	30	-	42	-	28
0310	-0.55	-	17	-	67	-	16
0311	-0.66	-	10	-	80	-	10
0312	-0.53	-	6	-	87	-	7
0408	+0.17	-	22	+	165	-	43
0409	+0.26	-	37	+	212	-	75
0410	-0.14	-	76	+	143	-	167
0411	-0.62	-	11	-	60	-	29
0412	-0.69	-	4	-	82	-	14
0508	+0.39	-	0	+	126	-	26
0509	+0.91	+	0	+	132	-	32
0510	+0.72	+	3	+	155	-	58
0511	-0.12	+	29	+	190	-	319
0512	-0.60	+	5	-	71	-	34
0608	+0.36	+	9	+	118	-	27
0609	+0.95	+	11	+	121	-	32
0610	+1.05	+	17	+	131	-	48
0611	+0.37	+	57	+	192	-	149
0612	-0.37	+	38	-	45	-	93
0708	+0.24	+	19	+	121	-	40
0709	+0.66	+	22	+	125	-	47
0710	+0.78	+	33	+	136	-	69
0711	+0.36	+	88	+	192	-	180
0712	-0.24	+	90	-	6	-	184
0808	+0.12	+	38	+	143	-	81
0809	+0.32	+	46	+	153	-	99
0810	+0.33	+	76	+	189	-	165
0811	+0.05	+	662	+	903	-	1465
0812	-0.27	+	76	-	3	-	173
0908	+0.02	+	163	+	375	-	438
0909	+0.04	+	340	+	685	-	925
0910	-0.06	+	327	+	481	-	908
0911	-0.26	+	86	+	61	-	247
0912	-0.31	+	44	-	12	-	132

tabulation. These come from an extra step done only at 50 cb, and show no change to the number of decimal places used in the table. Another implication of Table 4 is that there is a considerable tendency for V - and H -type increments to balance each other. This is evident in the smallness of the difference between the average absolute values of the initial and final approximate divergence compared with the average absolute values of the H - and V -type increments separately.

The last line of the table gives absolute average values over a set of 44 points, chosen for having the 10 extreme values of one or more of the quantities $D^{(0)}$, $\Delta^{(7)}D$, $\Delta^{(6)}D$, $\Delta^{(7)}D - \Delta^{(4)}D$,

and $\Delta^{(6)}D - \Delta^{(2)}D$ (collection A). As far as average absolute values can show it, this selection has the properties pointed out for the 20 points whose individual values are given in Table 4.

Subsequent tables summarize the results in various ways, to bring out more clearly characteristics such as those noted from Table 4. Table 5 is an organization of average values to show the coalescence of the iterative cycles. Two sets of points are considered, one a subset of the other. The larger consists of 72 points encompassing the central part of the region. The subset consists of the 44 points whose basis of choice is partly described in connection with their introduction in Table 4. Actually, it takes 22 points to represent the 10 extreme values of each of the 5 criterion quantities, but with each point that qualified there was included in the set its so-called 'counterpart point.' The counterpart point is merely the grid point in the symmetrical position

with respect to the longitudinal axis of the system.

Since inclusion of counterpart points in the subset dilutes the averages, the table shows in parentheses the average of the 10 extreme values for the criterion quantities.

The divergence contribution of the nondivergent-wind advection of vorticity is 0.5×10^{-6} sec⁻¹ for collection A, with an extreme of 1.5×10^{-6} sec⁻¹. The next 3 items of the table show the increments due to each of the cycles identified in Table 2. The first cycle, containing only an *H*-type increment, has an average contribution of 0.17×10^{-6} sec⁻¹. The difference between it and the contribution of the second cycle is about as large, owing primarily to the presence of a *V*-type correction for the first time. The third cycle replaces the second cycle by a value differing by 0.01×10^{-6} sec⁻¹.

This already indicates a rapid coalescence, in

TABLE 4. Successive Approximate Increments of Divergence on 50-cb Surface
Units are 10^{-6} per second.

Grid Point	$D^{(0)}$	$\Delta^{(1)}D$	$\Delta^{(2)}D$	$\Delta^{(3)}D$	$\Delta^{(4)}D$	$\Delta^{(5)}D$	$\Delta^{(6)}D$	$\Delta^{(7)}D$	$D^{(7)}$
0604	-0.65	-0.40	+0.51	-0.41	-0.41	+0.48	-0.41	-0.41	-0.58
0605	-1.54	-0.44	+0.27	-0.45	-0.44	+0.29	-0.44	-0.44	-1.68
0606	-1.30	+0.03	-0.21	+0.08	+0.08	-0.20	+0.08	+0.08	-1.42
0609	+1.30	-0.03	+0.32	-0.10	-0.12	+0.32	-0.11	-0.11	+1.50
0610	+1.54	+0.44	-0.18	+0.49	+0.47	-0.22	+0.46	+0.46	+1.78
0611	+0.65	+0.40	-0.56	+0.43	+0.43	-0.53	+0.43	+0.43	+0.55
0704	-0.65	-0.48	+0.49	-0.49	-0.49	+0.47	-0.48	-0.48	-0.67
0711	+0.65	+0.49	-0.51	+0.54	+0.53	-0.50	+0.52	+0.52	+0.67
0804	-0.21	-0.43	+0.33	-0.42	-0.40	+0.30	-0.40	-0.40	-0.31
0805	-0.55	-0.39	+0.23	-0.32	-0.32	+0.25	-0.32	-0.32	-0.62
0810	+0.55	+0.39	-0.04	+0.36	+0.33	-0.05	+0.33	+0.33	+0.82
0811	+0.21	+0.43	-0.36	+0.47	+0.46	-0.35	+0.45	+0.45	+0.32
0904	+0.23	-0.28	+0.10	-0.26	-0.24	+0.06	-0.25	-0.25	+0.04
0905	-0.03	-0.46	+0.29	-0.37	-0.36	+0.29	-0.36	-0.36	-0.10
0910	+0.03	+0.46	-0.19	+0.46	+0.43	-0.19	+0.41	+0.41	+0.25
0911	-0.23	+0.28	-0.16	+0.30	+0.29	-0.13	+0.29	+0.28	-0.08
1005	+0.33	-0.30	+0.10	-0.25	-0.22	+0.05	-0.22	-0.22	+0.15
1006	+0.15	-0.31	+0.22	-0.21	-0.21	+0.21	-0.21	-0.21	+0.15
1009	-0.15	+0.32	-0.10	+0.31	+0.29	-0.10	+0.27	+0.28	+0.02
1010	-0.33	+0.30	-0.14	+0.32	+0.31	-0.11	+0.30	+0.29	-0.15
Average absolute value for above points	0.56	0.35	0.27	0.35	0.34	0.255	0.34	0.34	0.59
Average absolute value for points of collection A*	0.67	0.26	0.20	0.26	0.25	0.19	0.25	0.25	0.72

* Collection A is a set of 44 points at which there occur the 10 extreme values of $D^{(0)}$, $\Delta^{(7)}D$, $\Delta^{(5)}D$, $\Delta^{(7)}D - \Delta^{(1)}D$, and/or $\Delta^{(5)}D - \Delta^{(2)}D$, plus the counterpart point of each (counterpart points are pairs located symmetrically with respect to grid $x = 0$). See text for further explanation.

the sense of the Cauchy criterion for approach to a limiting value. As far as numerical experimentation can indicate, the numerical process involved here has a limit and approaches it rapidly.

The third grouping of items in Table 5 brings out again the existence of a considerable amount of negative correlation between H - and V -type increments. The average absolute value of net change of divergence from first approximation $D^{(0)}$ to the accepted final approximation $D^{(7)}$ is the same thing as the average absolute value of the sum $\Delta^{(5)}D$ plus $\Delta^{(7)}D$. Cancellation between terminal V - and H -type terms is evident in the fact that this item in line 5 of the table is more than 3 times as large as the average of the sum of the absolute values. This characteristic shows up even more in a comparison of the same statistics formed for grid points having particularly large values of the divergence-linked increments of divergence.

The pronounced tendency for $\Delta^{(7)}D$ and $\Delta^{(5)}D$ to cancel each other is shown pictorially in Figure 12. On the left are superimposed analyses of computed vertical and horizontal divergence-linked increments of divergence at 50 cb. Isolines of $\Delta^{(5)}D$ are drawn as solid lines for unit intervals, with dashed intermediate lines. The solid lines interrupted by dots are isolines of $\Delta^{(7)}D$ at unit intervals, with dotted intermediate lines. The patterns are clearly different in detail, but work out to give a striking interference pattern which reduces the net divergence-linked contribution to a much smaller fraction of $D^{(0)}$ than if there were mutual reinforcement between the V - and H -type divergence contributions.

The diagram on the right in Figure 12 shows the distribution on the 50-cb surface of the measure of interference used in Table 5, namely the ratio

$$|\Delta^{(5)}D + \Delta^{(7)}D| : |\Delta^{(5)}D| + |\Delta^{(7)}D|.$$

It is plain that the areas in which the individual contributions are of greatest magnitude are precisely those in which the ratio is less than $\frac{1}{2}$. There is marked coherence of the regions of cancellation. The sets of points at which the ratio is large, indicating either reinforcement or that one of the contributions is predominant, are organized into contiguous groups to a lesser extent.

Referring again to Table 5, the last two

TABLE 5. Coalescence of Increments of Divergence in Iterative Procedure for 50 cb Units are 10^{-6} per second.

Increment	Average Absolute Value*	Extreme Value†	Average Absolute Value†
$D^{(0)}$	0.49	± 1.54	0.67 (1.08)
$D^{(1)} - D^{(0)}$	0.17	+0.49	0.26
$D^{(4)} - D^{(1)}$	0.15	-0.52	0.20
$D^{(7)} - D^{(4)}$	0.01	-0.05	0.02
$D^{(7)} - D^{(0)}$	0.10	+0.29	0.11
$ \Delta^{(5)}D + \Delta^{(7)}D $	0.32	+1.02	0.44
$\Delta^{(1)}D$	0.17	+0.49	0.26
$\Delta^{(4)}D - \Delta^{(1)}D$	0.03	-0.14	0.04
$\Delta^{(7)}D - \Delta^{(4)}D$	0.004	-0.02	0.005 (0.017)
$\Delta^{(2)}D$	0.16	-0.56	0.20
$\Delta^{(5)}D - \Delta^{(2)}D$	0.01	-0.05	0.02 (0.03)
$\Delta^{(8)}D - \Delta^{(5)}D$	0.001	-0.01	0.001 (0.004)

* Values in these columns are obtained for the 72 points in rows 5-10, columns 2-13 (collection B).

† Values in this column are for the collection A described in Table 4.

groupings of divergence increments are arranged to show the convergence of the H - and V -type increments individually. The first increments of each type are seen to average very much the same in absolute value for collection B, and are not very different for the smaller collection A. The table shows that iterative re-evaluation gives rapidly decreasing contributions of each kind. The indication is that three evaluations of each kind of divergence-linked effect from the vorticity equation give a value which is uncertain in the third decimal place times 10^{-6} per second. For the pressure surfaces above 50 cb at which the process was carried only through two evaluations of $\Delta^{(7)}D$ and one evaluation of $\Delta^{(5)}D$, the terminal value of divergence would probably be uncertain to an estimated maximum extent of 5×10^{-7} per second. At surfaces lower than 50 cb, at which the process was terminated with the third H -type increment and just before the third V -type, the final value is estimated to be good to 1×10^{-7} , owing to the truncation of the process. This estimate is made on the basis of the data from 50 cb illustrated in Table 5 together with knowledge from other study of the data. That is, it is found that the increments in the iteration are smaller throughout all cycles of iteration at the lower levels (where the model

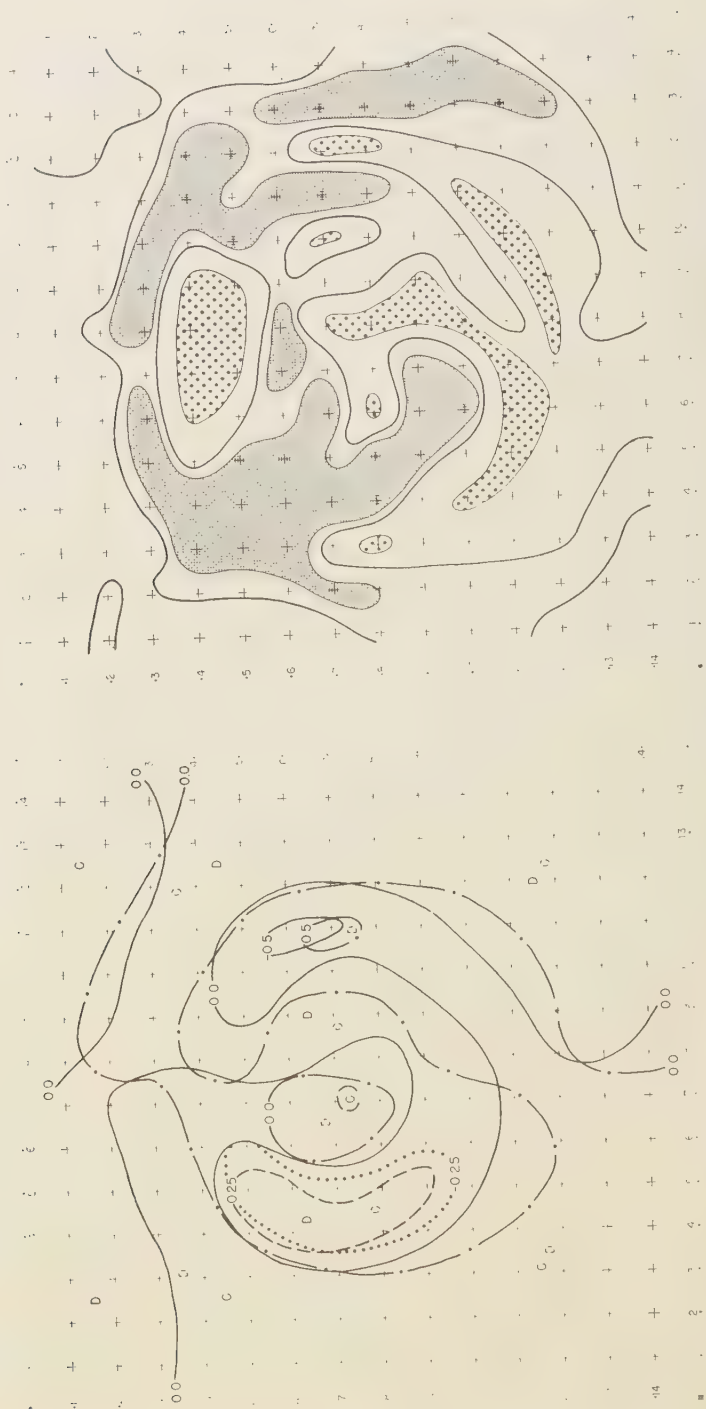


Fig. 12. *Left:* Superimposed scalar analyses of 50-cb divergence increments $\Delta^{(s)}D$ and $\Delta^{(v)}D$. Labels in units of 10^{-4} . *Right:* Scalar analysis of 50-cb distribution of ratio of $|D^{(v)} - D^{(s)}|$ to $|\Delta^{(s)}D| + |\Delta^{(v)}D|$. Gray area, 25 per cent or less; stippled area, 90 per cent or more.

TABLE 6. Summary of Successive Approximate Increments of Divergence at Upper Four Pressure Steps
Points of collection B, units 10^{-5} per second.

Values	$D^{(0)}$	$\Delta^{(1)}D$	$\Delta^{(2)}D$	$\Delta^{(3)}D$	$\Delta^{(4)}D$	$\Delta^{(5)}D$	$\Delta^{(6)}D$	$\Delta^{(7)}D$	$D^{(7)}$	$\Delta^{(8)}D$	$D^{(8)}$
60 cb											
Average											
absolute	0.30	0.10	0.08	0.09	0.09	0.08	0.09	0.09	0.30
extreme	± 0.85	± 0.29	-0.34	+0.29	+0.27	-0.21	+0.27	+0.27	+1.02
50 cb											
Average											
absolute	0.49	0.17	0.16	0.17	0.17	0.15	0.17	0.17	0.51	0.152	0.510
extreme	± 1.54	± 0.49	-0.56	+0.54	+0.53	-0.53	+0.52	+0.52	+1.78	-0.53	+1.78
40 cb											
Average											
absolute	0.76	0.31	0.30	0.32	0.31	0.29	0.30	0.30	0.78
extreme	± 2.49	± 0.92	-1.30	+1.04	+1.04	-1.22	+1.00	+0.99	+2.86
30 cb											
Average											
absolute	1.04	0.45	0.62	0.49	0.50
extreme	± 3.52	± 1.44	-3.02	+1.70	+1.75

and is weaker). The subsequent discussion will give further data on the results at pressure surfaces other than 50 cb.

The interpretation of such data as evidence that the terminal value of divergence is close to the true value required by the vorticity and continuity equations for the postulated model is purely extrapolative and empirical. There is no way to be sure that the results of this numerical experiment truly reveal the ultimate behavior of the iterative process. As a practical matter, however, the experimental results appear to be a satisfactory approximation to the solution of the problem undertaken.

Features of computed results at other pressure surfaces. For the purpose of inspection of other properties of the numerical process and the model, data of the type presented above for 50 cb are given in Tables 6 and 7 for other pressure steps. The point collection B consisting of the 72 grid

points in rows 5 through 10, columns 2 through 13, is the basis of these reduced data. Table 6 gives a summary for the four pressure values 60, 50, 40, and 30 cb, of the quantities tabulated for 50 cb in Table 4.

The following characteristics may be noted. The quantity $D^{(0)}$, measuring advection of absolute vorticity by the nondivergent wind, increases in the ratio of 1.5 to 1 for each 10-cb step. The increments produced by iterative computation, as exemplified by $\Delta^{(6)}D$ and $\Delta^{(7)}D$, are appreciable fractions of $D^{(0)}$. Both H - and V -type increments double with each 10-cb step upward.

In consideration of the 50-cb increments in Table 5, it is probable that termination of 40-cb computations after three cycles yields a value of $10^5 D^{(7)}$ that is uncertain in the second decimal place. At 60 cb and lower, $D^{(7)}$ was found by inspection of the results summarized in Table 6

TABLE 7. Average Absolute Value of Final V -Type and H -Type Increments of Divergence and Their Sums* (Relative to $D^{(0)}$) and the Cancellation Measure

p , cb	Average Fraction of $ D^{(0)} $			$ D^{(7)} - D^{(0)} $	Number of Grid Points
	$ \Delta^{(5)}D $	$ \Delta^{(7)}D $	$ D^{(7)} - D^{(0)} $	$ \Delta^{(5)}D + \Delta^{(7)}D $	
60	0.24	0.26	0.21	0.54	60
50	0.38	0.42	0.26	0.44	64
40	0.48	0.57	0.29	0.38	64

* Computed for points of collection B having $|D^{(0)}| \geq 0.10 \times 10^{-5}$.

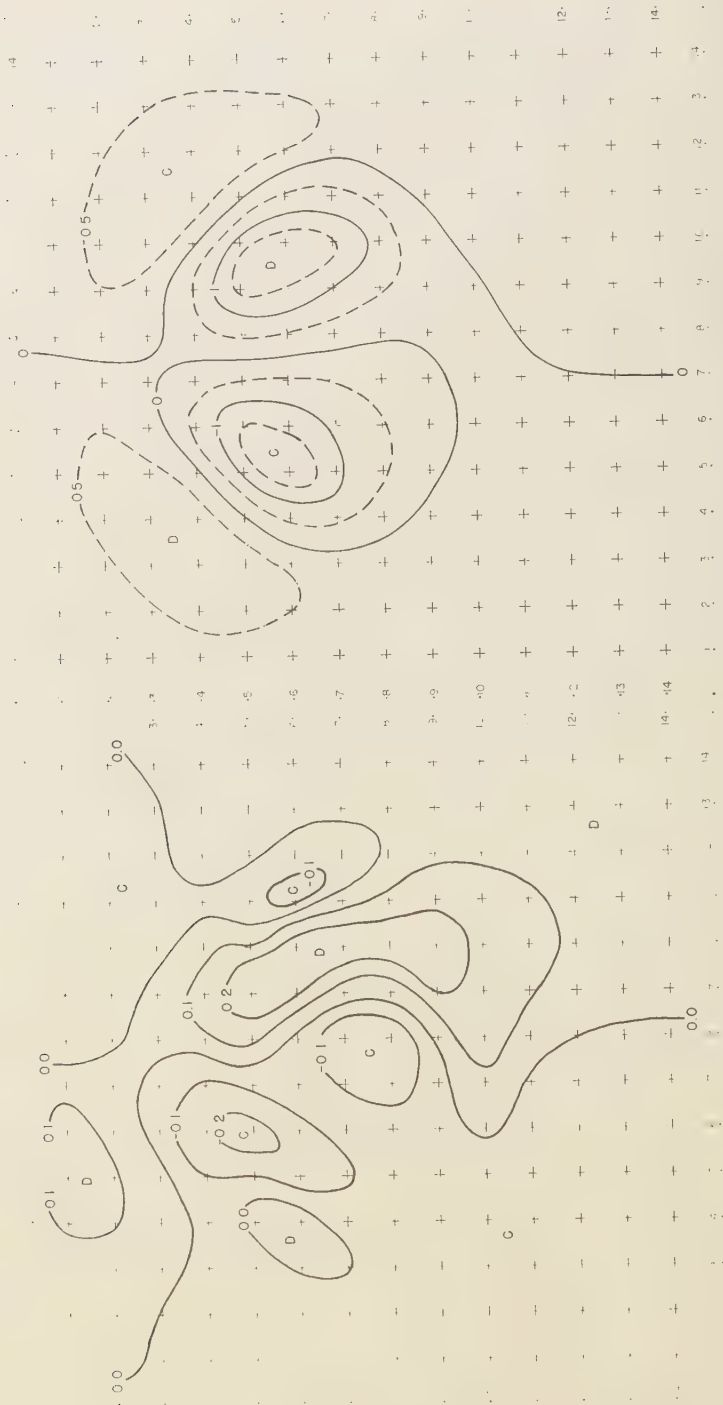


Fig. 13. Net change of divergence approximations for 50-cb³ surface, $D^{(1)}_{10} - D^{(0)}_{10}$, units 10^{-5} sec^{-1} . Fig. 14. Divergence approximation $D^{(7)}$ for 50 cb, units 10^{-5} sec^{-1} .

have the sought-for accuracy of 5×10^{-8} , as far as truncation of the over-all iteration is concerned. Early truncation at 40 cb, and the use of just two cycles at 30 cb, gives results that are amply accurate for their purpose. These being the last pressure steps, their role in error propagation is of no concern; they need only have enough accuracy to round out the experimental results.

The respective contribution of vertical and horizontal divergence-linked effects at all pressure surfaces is quite large. Each reaches the apparent final value to within 1×10^{-7} in one evaluation. Each type of correction taken alone, with random reinforcement from the other type, is large enough that the pattern of divergence based only upon nondivergent-wind advection of absolute vorticity could be radically altered by application of the full vorticity equation. However, systematic balancing between U - and V -type increments of divergence at all pressure surfaces, similar to that already pointed out for 50 cb, leads to the result that the final

divergence patterns are very similar to the patterns of $D^{(0)}$.

Table 7 gives the ratios of the final increments of each type to the zero-order approximation. This is essentially a comparison of the respective increments to the final divergence. The increments increase steadily with height in relation to the total divergence. Cancellations produce the result that the two kinds of increments put together have less effect than either type by itself. This is seen by comparison of the first two columns with the third. The extent of the cancellation effect in terms of the measure previously considered in Figure 12 is shown in the fourth column of Table 7. In these terms the cancellation is seen to become increasingly effective with height.

It is appropriate to comment at this point upon the relative magnitudes of vorticity and divergence in the atmospheric systems represented by the computed system. The average absolute value of the divergence at 50 cb is about $0.5 \times 10^{-6} \text{ sec}^{-1}$ for collection B. The correspond-

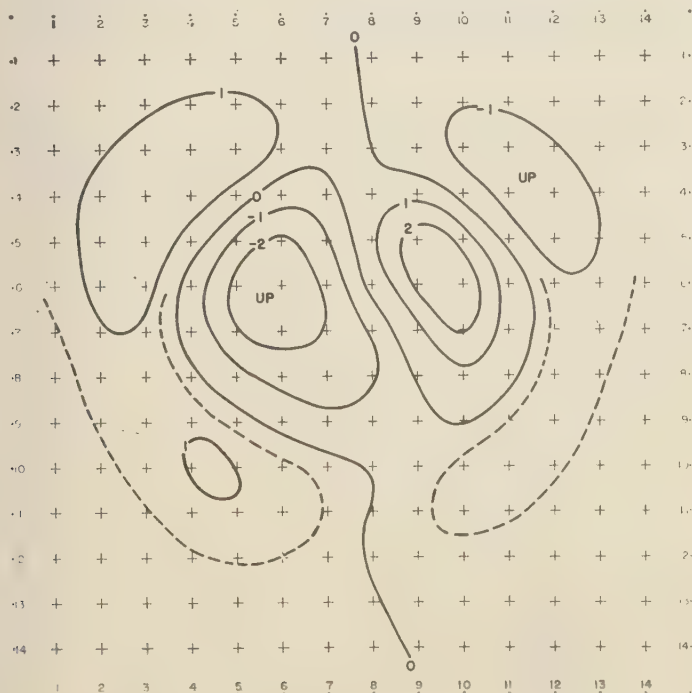


Fig. 15. Final approximation to ω for 50 cb, units $10^{-4} \text{ cb sec}^{-1}$.

cb

cb

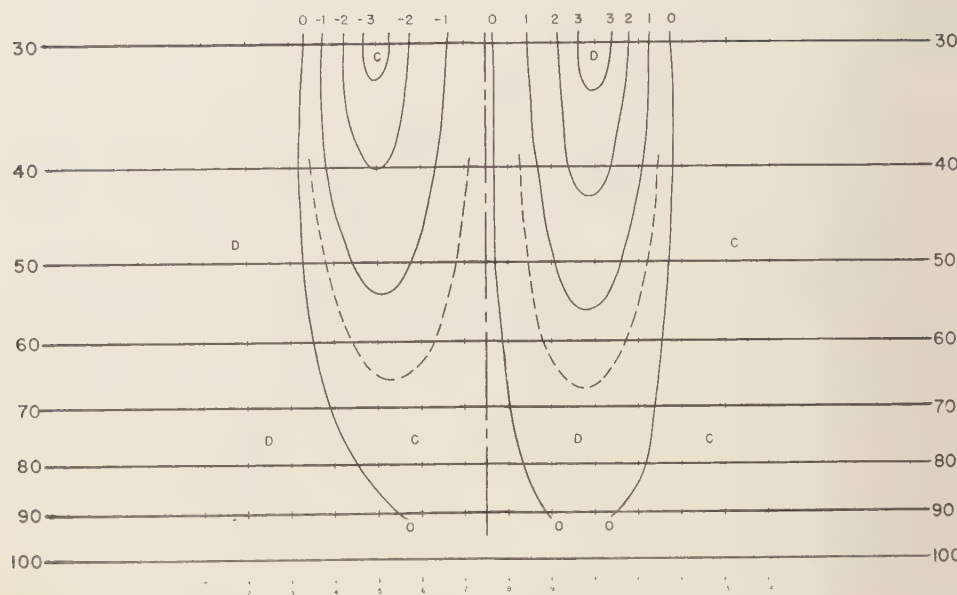


Fig. 16. West-east vertical cross section of divergence, units 10^{-5} sec^{-1} . Line of cross section is 150 km north of center of model vortex. Numbered crosses indicate columns on the grid.

cb

cb

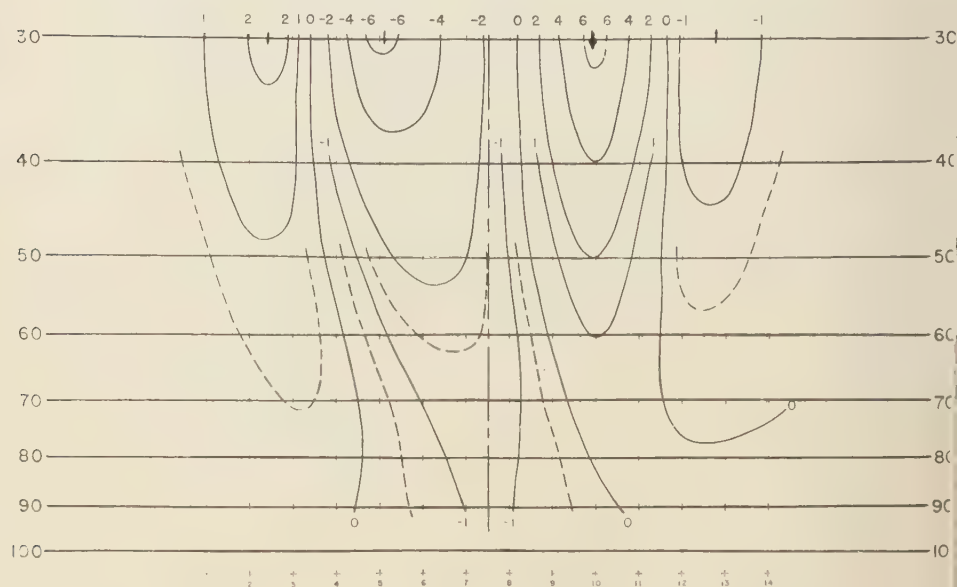


Fig. 17. West-east vertical cross section of vertical motion, units $10^{-4} \text{ cb sec}^{-1}$.

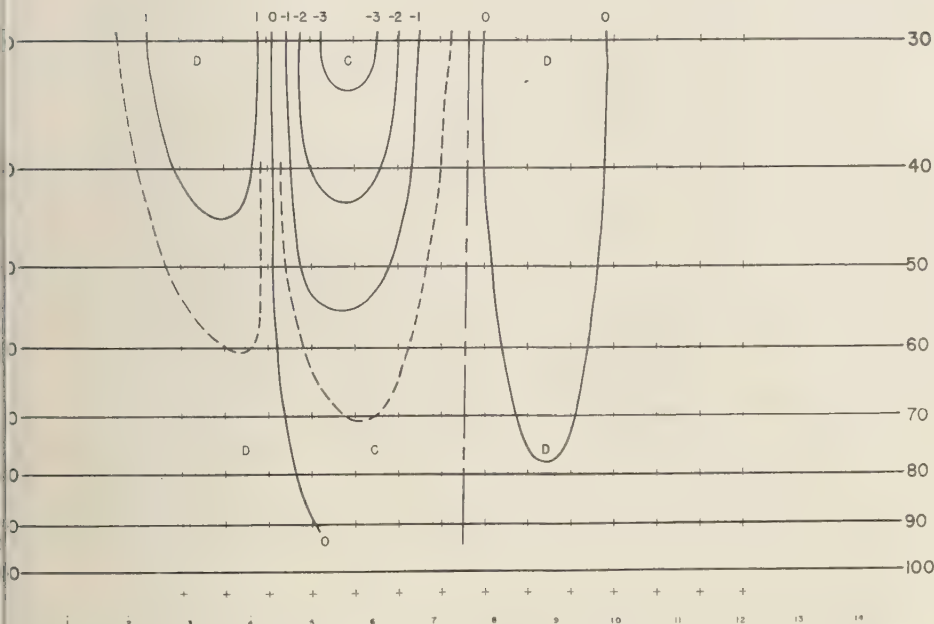


Fig. 18. Northwest-southeast vertical cross section of final approximate divergence, units 10^{-5} sec^{-1} . Numbered dots are grid points.

g mean relative vorticity is $3.7 \times 10^{-5} \text{ sec}^{-1}$. The respective maxima of divergence and vorticity are 1.8×10^{-5} and $14 \times 10^{-5} \text{ sec}^{-1}$. The quantitative comparisons for other pressure surfaces are about the same. For this wind field of moderate intensity, the average values and extreme values of relative vorticity are larger by an order of magnitude. These estimates are an important addition to the information available about divergence and vorticity in atmospheric systems. The exactness of method by which the estimates are obtained in this investigation is unique. They are not at all subject to the shortcomings of uncertain observations, gross finite-difference approximations, neglect of unobservables. To be sure, the direct approach of obtaining estimates by use of real observations exchanges these disadvantages for the advantage that no one can question the physical possibility of the collective data. One is free to raise such an objection to this modeling approach. In any case, this deduction of the divergence from the vorticity of a realistic wind field gives a complete motion system with

unshakable status as to internal consistency and considerable strength as to meteorological significance.

Pictorial representation of computed cold-low structure. Figures 13 through 15 show plan views of the 50-cb quantities whose distributions are of interest to the study of structure of the postulated type of atmospheric system. Figure 13 shows the pattern of change from the divergence which balances only the nondivergent-wind advection of absolute vorticity to the divergence required to balance all of the vorticity equation (with zero vorticity tendency in both cases). It may be noted that this pattern of H -plus V -type increments is more broken up than the corresponding $D^{(0)}$ pattern (Fig. 11), which in turn is about half the scale of the relative vorticity field. There is also considerably less symmetry in the pattern of computed increments of divergence.

Figures 14 and 15 show the horizontal distributions of divergence and vertical motion at 50 cb. These scalar analyses give essentially the accepted final values in the iterative process.

cb

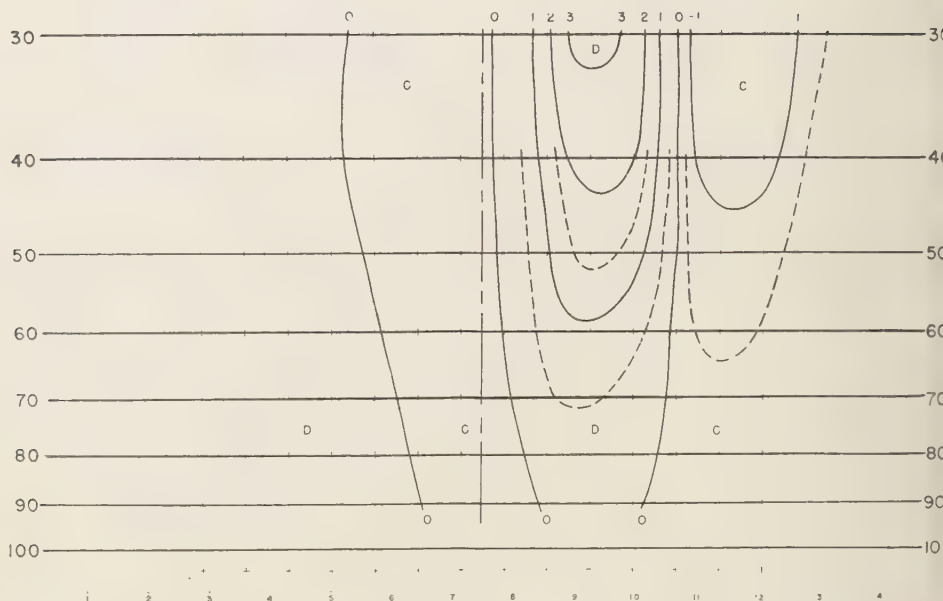


Fig. 19. Same as Figure 18, for southwest-northeast diagonal. Numbered dots are grid points.

The large element of $D^{(0)}$ pattern that remains upon iteration to the practical end is evident in a comparison of Figures 11 and 14. This comparison makes further interpretation of the characteristic balance between horizontal and vertical divergence-linked terms in the vorticity equation which has been brought out in foregoing statistics on behavior of the successive approximations.

These patterns for 50 cb are so representative of those for all pressure surfaces computed that there is no point in showing any others here. Another publication has the pictures for the pressure surfaces below 50 cb [Asplund, 1959].

Vertical cross sections of the final divergence and vertical motion from the lower boundary to the highest computed level are Figures 16 through 19. The lack of compensation along the vertical, and the placement of divergence and convergence centers, have emerged as consequences of previously discussed characteristics of the model and the deductive numerical process.

Several aspects of the deduced structure of the total system, complete with divergence and vertical motion, are disturbing. They suggest a

need for more work by this and other methods of study of the structure of atmospheric systems. On the question of the structure of cutoff high-level cyclones in particular, further application of the scientific method calls for looking carefully at case studies to see whether the computed results are reflected in the behavior of systems on weather maps.

Other features of computed results. The present formulation facilitates direct examination of the characteristics of the vorticity and divergence equations. Certain of these characteristics have been a source of interest in connection with numerical prediction formulation. Once the forecasts were refined as far as possible by empirical patching and engineering, it was found that refinement of models employing these derived equations probably have no great operational advantages over the primitive equations of motion. From the point of view of understanding the workings of the atmosphere, however, there are still good reasons for the study of the full vorticity and divergence equations.

One product of the present treatment of

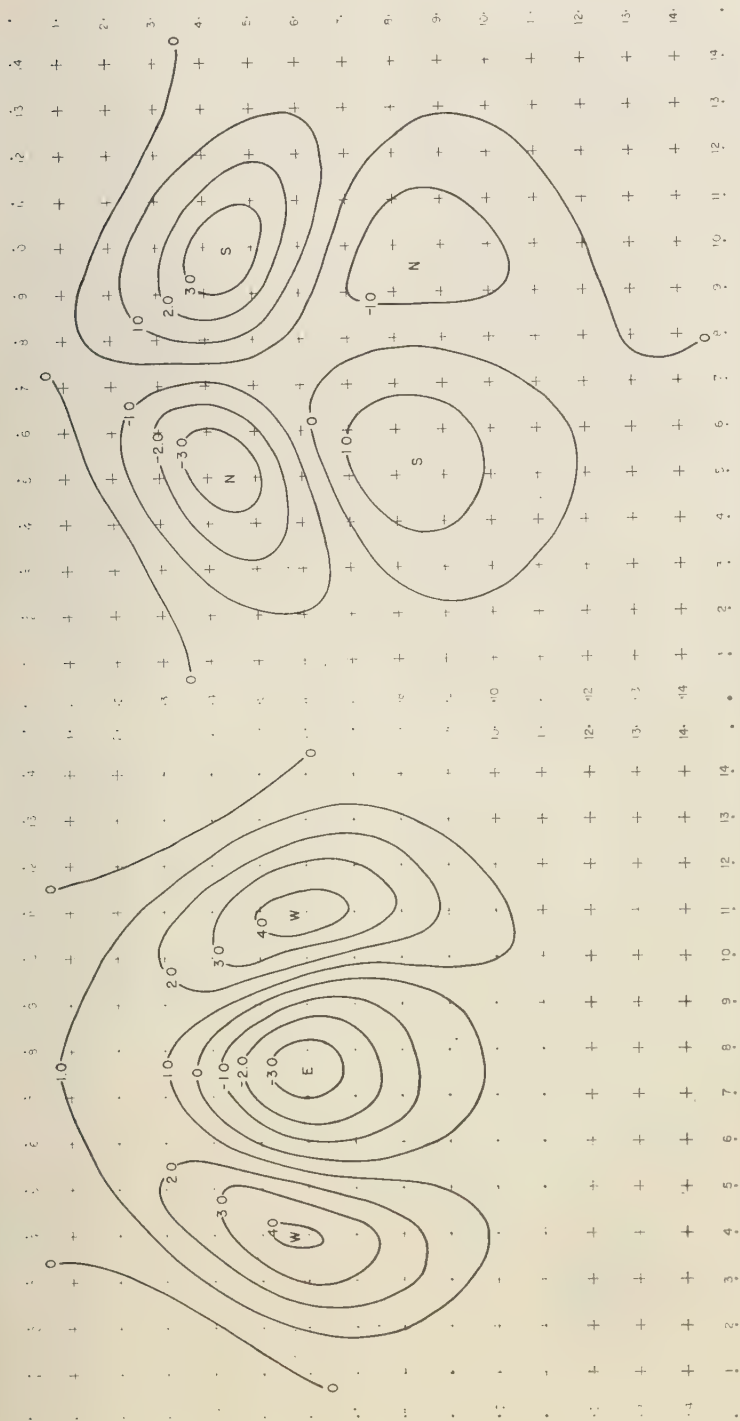


Fig. 20. Divergent wind on 50-cb surface, meters per second. *Left*: component in x direction. *Right*: y component.

equations of motion is explicit information on the divergent part of the wind field. Its effects in observed synoptic situations have been given attention by a few investigators (see, for example, an article by J. Knox now in press for the *Journal of Meteorology*). Here we are dealing with the wind field of a model. As was discussed above in another respect, a comparison of the observational basis against attempted simulation of ideal computing conditions finds advantages and disadvantages on both sides.

In brief mention of the computed divergent wind for the present model, the individual components in the grid x and y directions are shown in Figure 20. As is to be expected by the effect of the boundary conditions in restricting this component of the total wind field so as to contain only the pure divergent wind, the resultant pattern is closely and simply related to the divergence distribution.

It may be noted by reference to Figure 3 that the streamline field for the total model shows little effect of the presence of the divergent wind. This is a common property of the kinematics of atmospheric flow. The magnitudes of the respective rotational and divergent parts are not as disparate as one may intuitively think. As an extreme, we may note that at point 0511 on the 50-cb surface the divergent wind vector has a magnitude of 5 m sec^{-1} , nearly 30 per cent of the rotational part. On the average over 28 points at which it is relatively great, the divergent wind is 2.6 m sec^{-1} . The nondivergent component averages 21.9 m sec^{-1} over the same region, and the total wind averages 22.2 m sec^{-1} .

The results presented in earlier sections have shown explicitly the effect of the divergent wind as a part of the vorticity equation. We have seen that the $\Delta^{(2)}D$ term in question is not insignificant by itself, but is on the whole nearly canceled out for this model of zero vorticity tendency by the vertical divergence-linked effect $\Delta^{(v)}D$. Another property of the vorticity equation that may be readily studied with the output of the present numerical scheme is the cancellation within the $\Delta^{(v)}D$ term. This property has been of concern in connection with the possible use in prediction models of the vertical advection term without the tipping term [Arnason and Carstensen, 1959].

A study was made of this [Asplund, 1959] for a selection of points having large values of either

or both parts. The results did not show a clear-cut dominance by either part. There was moderate tendency for the parts to cancel, especially at the low levels. At 50 cb the tipping term averaged twice as large as the vertical advection term, and there were 11 occurrences of opposing signs against 5 reinforcements.

CONCLUSION

The three-dimensional space distributions of divergence and vertical motion on the synoptic scale surely merit systematic study and classification in terms of types of wind-field organization. A successful trial has been made of a numerical process for use in such a systematic investigation of atmospheric systems. The numerical experimentation presented indicates that the proposed procedure of mathematical modeling and successive approximation is stable and convergent. The design of the method permits its use whenever the nondivergent wind field and vorticity tendency are known for some volume of troposphere, if a vertical motion distribution can be given for one surface within or bounding the region.

This experiment dealt with a type of atmospheric system that is particularly amenable to certain expedients. Partly but not solely for this reason, an attempt to view the results as data on the structure of high-level cyclones was inconclusive. If they are viewed as a component of interplay of hypothesis and observation, more of the latter is needed than is now on hand.

Emphasis has been placed on the distinctive merits of data generated by the method. It appears that the numerical process gives reliable values for the divergence based on balancing the complete vorticity equation, with incident determination of the divergent part of the wind field. Given this success of the iterative procedure for the choices of horizontal and vertical lattice dimensions, finite-difference approximations by five-point formulas, and criterion of termination of the iterative processes of the relaxation subroutine, there are no procedural obstacles to very great refinement of the method.

The computations for this experiment were performed on the IBM 650. In order to inspect the results at fairly frequent intervals, and because of the computer's limited storage capacity, each of the two types of increments within one cycle was read out to cards. The success of the

process now being established, it can be speeded up by elimination of input and output routines to some extent with the IBM 650 and much more with a digital computer of greater capacity. It is estimated that the whole program for all iterations on a single pressure surface could be performed on a single loading of an 8000-word memory.

Information on divergence, vertical motion, and divergent wind obtainable for any specified model nondivergent wind field can be put to an interesting use. The divergence equation [Sherman, 1953; Petterssen, 1953] can be used as a Poisson equation for determination of the geopotential. By means of the relaxation process, the known function in this Poisson equation may be varied systematically from the $f\zeta$ term corresponding to computing vorticity by geostrophic assumption) through the terms of the balance equation [Petterssen, 1953; Kuo, 1955; Bolin, 1955; Charney, 1955] and on to the complete equation lacking only the local time-derivative of divergence. Taken all together, such further experimentation should produce interesting insight into the properties of the balance and divergence equations.

Acknowledgments. The investigation reported in this paper began with a suggestion of Professor Seymour L. Hess. I wish to acknowledge the benefit of many suggestions and discussions with him. I am also indebted to Walter James Koss and Harold D. Orville for many helpful discussions and a great amount of assistance in performance of the research. Continuation of the research in its last year, including investigation of the divergence equation, has been a joint project with Mr. Koss alone.

This work was supported by the Geophysics Branch, Office of Naval Research, contract Nonr-300(00), project NR-082-0711.

REFERENCES

- Bolin, G., A case study of the field of large-scale vertical velocity and horizontal divergence, *MIT Tech. Rept. 16*, ONR contract N5ori-07804, 70 pp., 1955.
- Bolin, G., and L. P. Carstensen, The effects of vertical vorticity advection and turning of the vortex tubes in hemispheric forecasts with a two-level model, *Monthly Weather Rev.*, **87**, 119-127, 1959.
- Bjund, S. E., A numerical experiment in deduction of the divergence associated with a model of a steady cold cyclone, *Fla. State Univ. Tech. Rept. 14*, contract Nonr 1600-(00), 180 pp., 1959.
- Bolin, B., Numerical forecasting with the baroclinic model, *Tellus*, **7**, 27-49, 1955.
- Brunt, D., *Physical and Dynamical Meteorology*, Cambridge University Press, London, 428 pp., 1939.
- Charney, J., The use of the primitive equations of motion in numerical prediction, *Tellus*, **7**, 27-49, 1955.
- Charney, J., and A. Eliassen, A numerical method for predicting perturbations of the middle latitude westerlies, *Tellus*, **1**, 38-54, 1949.
- Deardorff, J. W., The spatial distribution of vertical motion within a mature cyclonic vortex, *Sci. Rept. 3*, AF contract 19(604)-314, University of Washington, 46 pp., 1957.
- Eliassen, A., The quasi-static equations of motion with pressure as independent variable, *Geophys. Publikasjoner*, **17** (3), 1949.
- Eliassen, A., *Lectures on physical weather prediction* (mimeographed lecture notes), U.C.L.A., 109 pp., 1955.
- Fleagle, R. G., The fields of temperature, pressure, and three-dimensional motion in selected weather situations, *J. Meteorol.*, **4**, 165-185, 1947.
- Haltiner, G. H., and F. L. Martin, *Dynamical and Physical Meteorology*, McGraw-Hill Book Co., New York, 470 pp., 1957.
- Kunz, K. S., *Numerical Analysis*, McGraw-Hill Book Co., New York, 381 pp., 1957.
- Kuo, Hsiao-Lan, On quasi-nondivergent prognostic equations and their integration, *MIT Sci. Rept. 2*, General Circulation Project, 42 pp., 1955.
- Palmén, E., On the origin and structure of high-level cyclones south of the maximum westerlies, *Tellus*, **1**, 22-31, 1949.
- Petterssen, S., On the relationship between vorticity, deformation and the configuration of the pressure field, *Tellus*, **5**, 231-237, 1953.
- Petterssen, S., *Weather Analysis and Forecasting*, vol. 1, *Motion and Motion Systems*, McGraw-Hill Book Co., New York, 428 pp., 1956.
- Petterssen, S., and P. A. Calabrese, On some weather influences due to warming of the air by the Great Lakes in winter, *J. Meteorol.*, **16**, 646-652, 1959.
- Phillips, H. B., *Vector Analysis*, John Wiley & Sons, New York, 236 pp., 1933.
- Sherman, L., On the scalar-vorticity and horizontal-divergence equations, *J. Meteorol.*, **9**, 359-366, 1952.
- Sherman, L., Estimate of the vertical velocity based on the vorticity equation, *J. Meteorol.*, **10**, 399-400, 1953.

(Manuscript received December 21, 1960.)

On the Tensor Form of Dispersion in Porous Media

JACOB BEAR¹

University of California, Berkeley, California

Abstract. The variance of the bivariate normal distribution, which approximately defines the concentration distribution resulting from a tracer point injection into a uniform field of flow in a porous medium, is a second-rank tensor. When a point injection is subjected to a sequence of uniform movements in various directions, the final concentration distribution can be obtained by a summation of the tensors corresponding to the various movements. The concentration distribution across a transition zone, which develops when an abrupt interface between two miscible fluids is subjected to a sequence of uniform movements, can be determined by integrating the result for a single point injection over the entire tracer region.

The property of isotropic porous media to disperse a tracer fluid is defined by the constant of dispersion which is shown to be a fourth-rank tensor. If the displacement is defined as a second-rank tensor, the variance of the distribution is obtained by the product of twice the constant of dispersion and this displacement tensor.

Introduction. Tracer particles injected at one point into the liquid flowing in a porous medium will arrive after a given interval of time at different places depending upon the tortuosity of the path they followed. This phenomenon is called dispersion. The phenomenon of dispersion, or miscible displacement, has received much attention in recent years because of its importance in many fields of engineering. Comprehensive literature surveys on this subject were presented by Rifai, Kaufman, and Todd [1956], Lau [1959], and by Scheidegger [1959].

The dispersion phenomenon depends upon the geometry of the pore-grain system of the porous medium. An exact description of the flow pattern in a porous medium obtained by solving the Navier-Stokes equations is impossible because of the extremely complex nature of the surface of the impermeable solid matrix which serves as a boundary to the flow.

These difficulties were circumvented when the actual porous medium was represented by a simplified hypothetical model which could be treated analytically. Many investigators have used a model consisting of a bundle of capillary tubes for this purpose. The various models that were studied differed in diameter, distribution, and mode of orientation of the tubes.

The randomness of the geometry of the pore-

grain system suggested the application of statistical methods for the description of the porous medium. Several such statistical models were tried. It was assumed that the movement of each particle consists of a large number of elementary steps. Results were obtained by studying the microdynamics of the flow, i.e., what happens in each step or in each microscopic flow channel.

A continuous frequency function with respect to the direction of an elementary channel in the porous medium was used by de Josselin de Jong [1958]. The frequency function was obtained by assuming that it was proportional to the discharge of all channels oriented in that direction. The spreading caused by the dispersion in isotropic media is greater in the direction of the mean flow than normal to it. A mathematical description for this observed phenomenon was given by de Josselin de Jong [1958]. By computing the probability that a tracer particle will reach a given point after passing through a large number of elementary channels in the porous medium, he found that the spreading of the tracer from a point injection in a uniform, two-dimensional field of flow takes approximately the form of a bivariate normal distribution. This resulting distribution depends upon a longitudinal and a transverse constant of dispersion of the porous medium. This distribution is assumed throughout the present work.

¹ On leave from Technion-Israel Institute of Technology, Haifa, Israel.

Several investigators approached the problem of a tracer distribution in a field of flow by first constructing the partial differential equation which describes the tracer distribution and then solving it under the specified boundary condition. The partial differential equations themselves were arrived at by various assumptions. *Rifai, Kaufman, and Todd* [1956], for example, derived the equation

$$\partial C / \partial t = D \partial^2 C / \partial x^2 - \partial / \partial x (u C) \quad (1)$$

for the one-dimensional case, and

$$\partial C / \partial t = D \Delta C - \text{div} (u C) \quad (2)$$

for the three-dimensional case, by assuming that the dispersion in a laminar flow through porous media is basically due to molecular diffusion and velocity convection. The difference between a longitudinal and a lateral constant of dispersion was ignored in the generalized equation. They also summarized other earlier investigations leading to similar partial differential equations. *Lau, Kaufman, and Todd* [1959] derived in a similar manner the general two-dimensional dispersion equation in the form

$$\partial C / \partial t = D_1 u \partial^2 C / \partial r^2 - u \partial C / \partial r \quad (3)$$

for the radial case. Because of symmetry, the lateral effect is cancelled.

The author [Bear, 1960], through a cell-balance analysis, arrived at the equation

$$\partial C / \partial t = D_1 |u| \partial^2 C / \partial x^2 - u_2 \partial C / \partial x \quad (4)$$

for one-dimensional flow, and

$$\frac{\partial C}{\partial t} = D_1 \left\{ |u| \frac{\partial^2 C}{\partial x^2} + |v| \frac{\partial^2 C}{\partial y^2} \right\} - u \frac{\partial C}{\partial x} - v \frac{\partial C}{\partial y} \quad (5)$$

for a two-dimensional field of flow. However, no account was taken in (5) of tracer interchange across the elementary cell boundaries due to lateral dispersion, and the equation was considered a rough approximation only.

Some other investigators believe that the diffusivity equation

$$\frac{\partial C}{\partial t} = \frac{\partial}{\partial x_i} D_{ik} \frac{\partial C}{\partial x_k} - \frac{\partial}{\partial x_i} (v_i C) \quad (f)$$

with $C(x_i)$ being the tracer concentration and D_{ik} the dispersion coefficient which depends in

an unknown fashion on the velocity component v_i , describes the dispersion of a tracer. The equation, again, is obtained by considering the dispersion to be a sum of two processes, namely, the diffusion process described by the first term on the right and the convection process described by the second term.

Rather than attempting to derive the general equation and solve it under various boundary conditions, a different approach is presented in the following paragraphs. The question of the general equation was purposely left out of the discussion at this stage, although it seems that some of the conclusions of the present paper should later be tied with the general equation of tracer motion.

The influence of molecular diffusion is omitted from the present work.

The purpose of this work is to study the relationship between the dispersive property of the porous medium as defined by a constant of dispersion, the displacement caused by a uniform field of flow, and the resulting distribution.

A point injection subjected to a sequence of movements. By using the normal frequency function obtained by *de Josselin de Jong* [1958], the concentration distribution of a point injection of a tracer of concentration C_0 around a point (ξ, η) which is displaced a distance $L = ut$ in the direction of the uniform, isotropic, two-dimensional field of flow from its original position (x_0, y_0) is given by

$$C(x, y; x_0, y_0) = \frac{C_0}{2\pi\sigma_x\sigma_y} \cdot \exp \left\{ -\frac{m^2}{2\sigma_x^2} - \frac{n^2}{2\sigma_y^2} \right\} \quad (6)$$

where

L = distance of mean displacement (see Fig. 1).

u = uniform velocity of flow.

t = time of flow.

σ_x, σ_y = standard deviations of the distribution in the x and y directions respectively.

m, n = the coordinates of the point (x, y) the coordinate system centered (ξ, η) . These are given by

$$m = x - (x_0 + L)$$

$$n = y - y_0$$

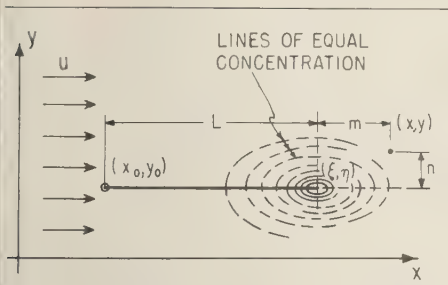


Fig. 1. Dispersion of a point injection displaced a distance L .

The standard deviations are defined by

$$\sigma_x = (2 D_I L)^{1/2}; \quad \sigma_y = (2 D_{II} L)^{1/2}$$

where D_I and D_{II} (dimensions, cm) are the longitudinal and transverse constants of dispersion of the porous media, respectively. Lau [1959], and others later, used the notation D_m to denote the longitudinal constant of dispersion and defined D_m (dimensions, cm) by the ratio between the coefficient of dispersion D (dimension, m^2/sec) and the average velocity of flow through the pores. The parameter D , which measures the extent of the longitudinal dispersion in the one-dimensional flow and correlates with both the physical characteristics of the porous media and the flow, is not used in this work. Instead, D_I and D_{II} (dimensions of both, cm), which depend only upon properties of the porous medium such as porosity, grain size, uniformity, and shape of grains, are used. It was found necessary to drop the index m and to change notation because of the special meaning that indices obtain in the following paragraphs.

Consider a case where at the end of a period t_1 the velocity, remaining horizontal, is suddenly changed from u_1 to u_2 and remains unchanged for a period t_2 . The mean displacement in the

second period is, therefore, $L_2 = u_2 t_2$. The tracer concentration at a typical point (x_1, y_1) (Fig. 2) at the end of the first period can be used as a new point source with concentration $C(x_1, y_1; x_0, y_0)$ obtained from (6). The concentration at a point (x, y) at the end of the second period resulting from this single point source is given by

$$C(x, y; x_1, y_1) = \frac{C(x_1, y_1; x_0, y_0)}{2\pi\sigma_x^{(2)}\sigma_y^{(2)}} \cdot \exp \left\{ -\frac{(x - x_1 - L_2)^2}{2\sigma_x^{(2)2}} - \frac{(y - y_1)^2}{2\sigma_y^{(2)2}} \right\} \quad (7)$$

where $\sigma_x^{(2)}$ and $\sigma_y^{(2)}$ are the longitudinal (i.e., in the direction of flow) and the lateral standard deviations of the concentration distribution.

The distribution of the tracer at the end of the second period is obtained by integrating over all the points (x_1, y_1) :

$$C(x, y) = \int_{-\infty}^{+\infty} dx_1 \int_{-\infty}^{+\infty} \left[\frac{C(x_1, y_1; x_0, y_0)}{2\pi\sigma_x^{(2)}\sigma_y^{(2)}} \cdot \exp \left\{ -\frac{(x - x_1 - L_2)^2}{2\sigma_x^{(2)2}} - \frac{(y - y_1)^2}{2\sigma_y^{(2)2}} \right\} \right] dy_1 \quad (8)$$

Introducing the value of $C(x_1, y_1)$ obtained from (6) with (x_1, y_1) replacing (x, y) and performing the integration yields the following result:

$$C(x, y) = \frac{C_0}{2\pi\bar{\sigma}_x\bar{\sigma}_y} \cdot \exp \left\{ -\frac{[x - (x_0 + L)]^2}{2\bar{\sigma}_x^2} - \frac{[y - y_0]^2}{2\bar{\sigma}_y^2} \right\} \quad (9)$$

where

$$L = L_1 + L_2 = u_1 t_1 + u_2 t_2$$

$$\bar{\sigma}_x^2 = \sigma_x^{(1)2} + \sigma_x^{(2)2} = 2 D_I L$$

$$\bar{\sigma}_y^2 = \sigma_y^{(1)2} + \sigma_y^{(2)2} = 2 D_{II} L$$

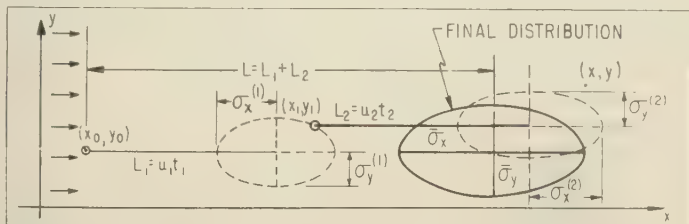


Fig. 2. Dispersion of a point injection displaced a distance $L_1 + L_2$.

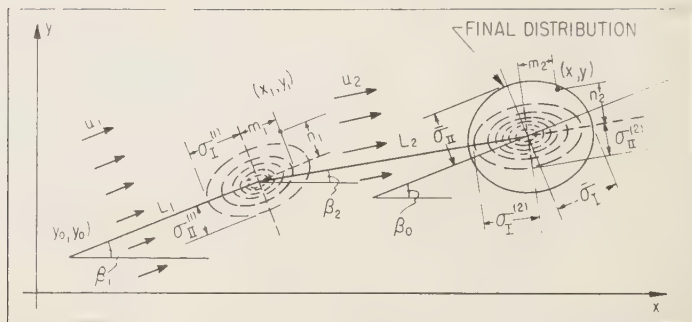


Fig. 3. Dispersion of a point injection when the uniform velocity field is changed.

Equation 9 is again a bivariate normal frequency function with mean at $(x_0 + L, y_0)$ and standard deviations $\bar{\sigma}_x$ and $\bar{\sigma}_y$. Consequently, for and variations of velocity, the final distribution depends on the total path as long as the uniform flow remains horizontal.

The same conclusion can be obtained for the one-dimensional case [Bear 1960] by solving (4). For example, with the boundary conditions

$$x = \pm \infty \quad t \geq 0 \quad \frac{\partial C}{\partial x} = 0$$

$$x = +\infty \quad t \geq 0 \quad C = 1$$

$$x = -\infty \quad t \geq 0 \quad C = 0$$

and initial conditions

$$t = 0 \quad -\infty < x < +\infty \quad C = 0$$

$$0 \leq x < +\infty \quad C = 1$$

one obtains the solution

$$C(x, t) = \frac{1}{2} \operatorname{erfc} \left\{ \frac{x - \int_0^t u(t) dt}{\left[4 D_I \int_0^t |u| dt \right]^{1/2}} \right\}$$

One should recall that approximately a normal distribution is obtained only after a large number

of elementary displacements in the pore-channel system is included. Therefore, whenever a period with constant flow is mentioned in this work it is considered to include a sufficiently large number of individual displacements to justify the assumption of a normal frequency function.

The problem can now be generalized to the case in which after a period t_1 both the velocity and the direction of the field of flow are changed (Fig. 3). The solution involves repeating the above integration procedure to obtain the desired tracer distribution. With the nomenclature of Figure 3, the distribution at the end of the first period is given by

$$C(x_1, y_1) = \frac{C(x_0, y_0)}{2\pi \sigma_{I1}^{(1)} \sigma_{II1}^{(1)}} \cdot \exp \left\{ -\frac{m_1^2}{2\sigma_{I1}^{(1)2}} - \frac{n_1^2}{2\sigma_{II1}^{(1)2}} \right\} \quad (10)$$

where (x_1, y_1) is an arbitrary point in the field

$$\sigma_{I1}^{(1)} = (2 D_I L_1)^{1/2}; \quad \sigma_{II1}^{(1)} = (2 D_{II} L_1)^{1/2}$$

$$m_1 = (y_1 - y_0) \sin \beta_1 + (x_1 - x_0) \cos \beta_1 - L$$

$$n_1 = (y_1 - y_0) \cos \beta_1 - (x_1 - x_0) \sin \beta_1$$

Consequently, at the end of the second period the distribution in the entire plane is given by

$$C(x, y) = \frac{C(x_0, y_0)}{(2\pi)^2 \sigma_{I1}^{(1)} \sigma_{II1}^{(1)} \sigma_{I2}^{(2)} \sigma_{II2}^{(2)}} \cdot \int_{-\infty}^{+\infty} dx_1 \int_{-\infty}^{+\infty} \left[\exp \left\{ -\frac{m_1^2}{2\sigma_{I1}^{(1)2}} - \frac{n_1^2}{2\sigma_{II1}^{(1)2}} - \frac{m_2^2}{2\sigma_{I2}^{(2)2}} - \frac{n_2^2}{2\sigma_{II2}^{(2)2}} \right\} \right] dy_1 \quad (11)$$

here

$$\sigma_{II}^{(2)} = (2 D_{II} L_2)^{1/2}; \quad \sigma_{II}^{(2)} = (2 D_{II} L_2)^{1/2}$$

$$m_2 = (y - y_1) \sin \beta_2$$

$$+ (x - x_1) \cos \beta_2 - L_2$$

$$n_2 = (y - y_1) \cos \beta_2 - (x - x_1) \sin \beta_2$$

The results obtained by a detailed integration (11) are given by

$$C(x, y) = \frac{C(x_0, y_0)}{2\pi\sigma_I\sigma_{II}} \cdot \exp \left\{ -\frac{[(y - \bar{y}) \sin \beta_0 + (x - \bar{x}) \cos \beta_0]^2}{2\sigma_I^2} - \frac{[(y - \bar{y}) \cos \beta_0 - (x - \bar{x}) \sin \beta_0]^2}{2\sigma_{II}^2} \right\} \quad (12)$$

here

$$\bar{x} = x_0 + L_1 \cos \beta_1 + L_2 \cos \beta_2$$

$$\bar{y} = y_0 + L_1 \sin \beta_1 + L_2 \sin \beta_2$$

$$\sigma_I^2, \sigma_{II}^2 = (D_I + D_{II})(L_1 + L_2)$$

$$\pm (D_I - D_{II})[(L_1 \cos 2\beta_1 + L_2 \cos 2\beta_2)^2 + (L_1 \sin 2\beta_1 + L_2 \sin 2\beta_2)^2]^{1/2}$$

$$\tan 2\beta_0 = (L_1 \sin 2\beta_1 + L_2 \sin 2\beta_2)/(L_1 \cos 2\beta_1 + L_2 \cos 2\beta_2) \quad (13)$$

Even for this simple case of two steps, performing the integration is a very lengthy and tedious procedure. The integration is almost impossible for a large number of steps. A different approach to the problem is, therefore, proposed in the following sections.

From (13) one can already notice the special manner in which the displacements L_i are

summed up. This indicates that the L_i 's might be considered to be second-rank tensors.

Before leaving the method of direct integration, another case will be discussed so that its result might later be used for comparison with results obtained by the proposed new approach.

Consider the case of an abrupt interface between two miscible fluids (Fig. 4).

The resulting concentration distribution after a sequence of uniform movements with velocities u_i prevailing for periods t_i may be obtained by adding the resulting distributions from point sources covering the half-plane of the tracer fluid region. Thus, integrating (9) for an interface which is perpendicular to the direction of the flow yields

$$C(x, y) = \int_{-\infty}^{+\infty} d\eta \int_{-\infty}^{\eta} \frac{1}{2\pi\sigma_x\sigma_y} \cdot \exp \left\{ -\frac{(x - \xi - L)^2}{2\sigma_x^2} - \frac{(y - \eta)^2}{2\sigma_y^2} \right\} d\xi = \frac{1}{2} \operatorname{erfc} \left\{ \frac{x - L}{(2\sigma_x^2)^{1/2}} \right\} \quad (14)$$

This normal distribution with mean at $x = L$ and standard deviation σ_x describes the tracer distribution in the transition zone which develops as a result of the motion.

The same procedure can be repeated when the interface makes an angle α with the direction of the flow. The result is given by

$$C(x, y) = \int_{-\infty}^{+\infty} d\eta \int_{-\infty}^{\eta \cot \alpha} \frac{1}{2\pi\sigma_x\sigma_y} \cdot \exp \left\{ -\frac{(x - \xi - L)^2}{2\sigma_x^2} - \frac{(y - \eta)^2}{2\sigma_y^2} \right\} d\xi = \frac{1}{2} \operatorname{erfc} \left\{ \frac{x - y \cot \alpha}{[2(\sigma_x^2 + \cot^2 \alpha \sigma_y^2)]^{1/2}} \right\} \quad (15)$$

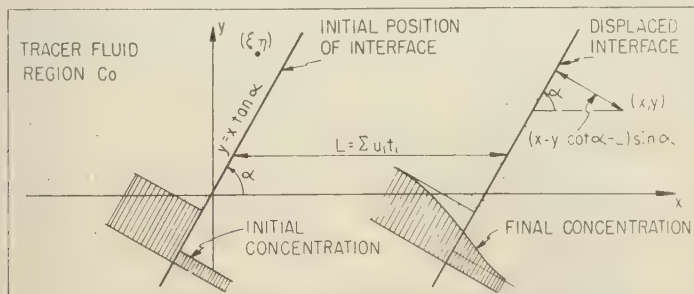


Fig. 4. Transition zone at an interface.

Hence

$$C(x, y) = \frac{1}{2} \operatorname{erfc} \left\{ \frac{x \sin \alpha - y \cos \alpha - L \sin \alpha}{[2(\sigma_x^2 \sin^2 \alpha + \sigma_y^2 \cos^2 \alpha)]^{1/2}} \right\} \quad (16)$$

Equation 16 describes a normal distribution perpendicular to the displaced interface.

The numerator in the argument of the complementary error function of (16) is the distance from the point (x, y) to the displaced interface. If the standard deviation of the distribution of (16) is denoted by σ_n , then

$$\sigma_n^2 = \sigma_x^2 \sin^2 \alpha + \sigma_y^2 \cos^2 \alpha \quad (17)$$

or

$$\sigma_n^2 = 2 D_\alpha L \quad (18)$$

where

$$D_\alpha = D_I \sin^2 \alpha + D_{II} \cos^2 \alpha \quad (19)$$

This defines a directional constant of dispersion associated with the direction α . It will later be shown that (17) describes an intrinsic property of the dispersion and that the variance of the distribution has a tensorial nature.

When the motion is parallel to the interface, i.e., $\alpha = 0$, equation 16 yields

$$C(x, y) = \frac{1}{2} \operatorname{erfc} \{ y / (2\sigma_y^2)^{1/2} \} \quad (20)$$

$$\sigma_y = (2 D_{II} L)^{1/2}$$

Thus, the lateral constant of dispersion determines the concentration distribution in the transition zone which develops when the interface between the two fluids is in the direction of the flow.

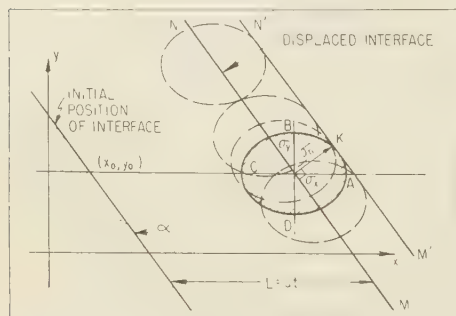


Fig. 5. The ellipse of tracer distribution (at the standard deviation) from a point injection.

The ellipse of dispersion. From (6) it follows that, after a period of uniform flow, lines of equal concentration resulting from a point injection of the tracer take the shape of ellipses centered at the displaced mean point and oriented with their major axes in the direction of the flow. The equation of an ellipse corresponding to a certain tracer concentration is obtained by letting the expression under the exponential sign be equal to a certain constant. The ellipse for which the constant is equal to $-\frac{1}{2}$ corresponds to the concentration at the standard deviation. Since the distribution is Gaussian in each direction, this ellipse uniquely defines the tracer concentration distribution. The radius vector from the center of the ellipse to any point on it gives the standard deviation in that direction. An ellipse is uniquely defined in the plane if its major and minor axes as well as the orientation of the major axis are known. If the major and minor axes are σ_x and σ_y parallel to the x and y axes, respectively, then the equation of this ellipse is

$$x^2/\sigma_x^2 + y^2/\sigma_y^2 = 1 \quad (21)$$

These ellipses were used to define the tracer distribution in Figures 1 to 3. The ellipse $ABCD$ in Figure 5 describes the distribution from the point injection at (x_0, y_0) . When an abrupt interface passes through the point (x_0, y_0) , each point along it as well as in the tracer fluid region give rise to such an ellipse. The resulting distribution which was previously determined in the form of (16), can be defined by its standard deviation perpendicular to the displaced interface (MN in Fig. 5). This standard deviation σ_n is obtained by drawing the tangent $M'N'$ to the ellipse $ABCD$ parallel to MN . The distance σ_n to this tangent can be shown to be equal to

$$\sigma_n = (\sigma_x^2 \sin^2 \alpha + \sigma_y^2 \cos^2 \alpha)^{1/2} \quad (22)$$

This result was previously obtained by direct integration in (17).

The tensor of variances. In any distribution the second moment can be used as a measure of the spread. If X is a random variable with second moment $E(X^2)$ and mean $\mu = E(X)$, then a number called the variance of X is defined [see Feller, 1957] by

$$\text{Var}(X) = E((X - \mu)^2) = E(X^2) - \mu^2 \quad (23)$$

Sometimes the term 'dispersion' is used as

onym for 'variance.' The standard deviation is defined as the positive square root of the variance and is generally used as a measure of dispersion rather than the variance itself. The relationship between the variance and the standard deviation is given, therefore, by

$$\text{Var}(X) = \sigma^2 \quad (24)$$

In general, whereas the mean point of a two-dimensional distribution corresponds to the center of gravity of the mass in the distribution (i.e., the mass of the tracer fluid in the present case), the variance, or the second central moment, corresponds to the moment of inertia of the tracer mass.

The moment of inertia of a mass distribution obtains different values when related to different axes passing through the center of gravity, or the mean point of the distribution. The bivariate normal distribution is uniquely defined by its mean point and the moments of inertia of its mass distribution with respect to two perpendicular axes through its mean point.

In view of the above considerations, the distribution resulting from a point injection of a tracer into a uniform field of flow could be described after a period t either by its standard deviation in the direction of the flow (σ_I) and that perpendicular to it (σ_{II}), with

$$\sigma_I^2 = 2 D_I L; \quad \sigma_{II}^2 = 2 D_{II} L \quad (25)$$

by the variances a_{11} and a_{22} defined by

$$a_{11} = \sigma_I^2; \quad a_{22} = \sigma_{II}^2 \quad (26)$$

or reasons given below, it is more convenient to use the variance as a measure of the dispersion of the tracer fluid.

If $C(x, y)$ is the tracer concentration at each point in the plane, three moments of inertia are defined with respect to the two perpendicular axes x and y :

Moment of inertia with respect to the y axis,

$$I_y = \int x^2 C(x, y) dx dy / \int C(x, y) dx dy$$

Moment of inertia with respect to the x axis,

$$I_x = \int y^2 C(x, y) dx dy / \int C(x, y) dx dy$$

Moment of inertia with respect to x and y ,

$$a_{xy} = \int xy C(x, y) dx dy / \int C(x, y) dx dy \quad (27)$$

when x and y are the principal axes, $a_{xy} = 0$.

The moment of inertia of a mass distribution is a second-rank tensor; therefore, in the graphical and analytical procedures presented in the following paragraphs, methods and operations of tensor algebra are introduced.

A physical quantity which is said to be a tensor of a certain rank is characterized by the way it behaves under a change in the coordinate system. Although formulas are available for this purpose, it is believed that a graphical representation of the variance describing the tracer distribution will lead to a better understanding of the problem under consideration.

A common graphical representation for a second-rank symmetric tensor is an ellipse. With a_{11} as major axis (in the direction of the flow) and a_{22} as minor axis (perpendicular to the direction of the flow), an ellipse $ABCD$ with center at the mean point of the distribution (0) can be constructed (Fig. 6). In this case a_{11} and a_{22} are the principal (maximum and minimum) moments of inertia with respect to the y and x axes, respectively. It should be noted that the axes of this ellipse are different from those of the ellipses described in Figures 1, 2, and 3, in which standard deviations are used. In Figure 6 and in subsequent figures, the ellipse of variances should be regarded as a diagram used for the explanation of certain operations and not as describing the actual tracer distribution.

With any point M on the ellipse and a radius vector a_α (Fig. 6) there is associated a certain plane p with normal n . The construction of the point M when a plane p and its normal n are given is as follows. The normal n intersects the circles $A'BC'D$ with radius a_{22} and $AB'CD'$ with radius a_{11} at E and F , respectively. From F and E lines parallel to the y and x axes, respectively, are drawn. The point M is at the intersection of these two lines.

a_α is the moment of inertia (of the tracer mass distribution) associated with the line p . When a_α is resolved into a_{nn} in the direction of n and a_{pn} in the direction of p , then a_{nn} is the moment of inertia with respect to p , and a_{pn} defines the

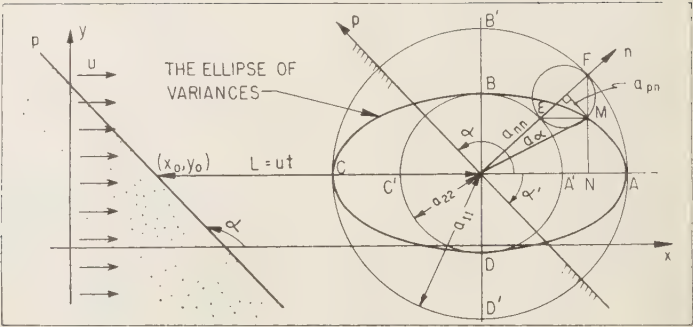


Fig. 6. The ellipse of variances.

moment of inertia with respect to p and n . These moments of inertia are defined by (27), with n and p replacing x and y , respectively. From the triangle OMN it can be shown that

$$a_{\alpha}^2 = a_{11}^2 \sin^2 \alpha + a_{22}^2 \cos^2 \alpha \tag{28}$$

$$a_{nn} = a_{11} \sin^2 \alpha + a_{22} \cos^2 \alpha \tag{29}$$

or, in another form,

$$a_{nn} = \frac{1}{2}(a_{11} + a_{22}) - \frac{1}{2}(a_{11} - a_{22}) \cos 2\alpha$$

and

$$a_{pn} = -\frac{1}{2}(a_{11} - a_{22}) \sin 2\alpha \tag{30}$$

From (28) it follows that

$$a_{nn} = \sigma_n^2 = \sigma_I^2 \sin^2 \alpha + \sigma_{II}^2 \cos^2 \alpha \tag{31}$$

With $\sigma_I = \sigma_x$ and $\sigma_{II} = \sigma_y$, σ_n is the same σ_n that was discussed previously and shown in Figure 5; it describes the actual tracer distribution. When the square root of a_{nn} , as obtained from Figure 6, is thus drawn in Figure 5 perpendicular to MN , which makes the same angle α with the positive

x axis, a point K is obtained. The locus of all points K for all the ellipses with centers on MN is the line $M'N'$ parallel to MN . This is the line of equal concentration at the standard deviation for a plane, making an angle α with the positive x axis, that has been displaced a distance L parallel to the positive x axis from its original position. This result is the same as the one previously obtained by direct integration.

It is impossible to compare the value of a_{pn} from (30) with a value obtained by direct integration, because it is eliminated when the ensemble of points making the tracer fluid region is considered. It has no physical meaning in the present case.

As the plane p turns around the point (x_0, y_0) the point M in Figure 6 traces an ellipse.

Consider, next, the case in which the velocity u makes an angle β with the positive x axis (Figure 7). Let λ be the angle between the normal to the plane and the positive x axis. From Figure 7,

$$\lambda = \beta + \varphi = \alpha - 90^\circ \tag{32}$$

$$a_{nn} = \frac{1}{2}(a_{11} + a_{22}) + \frac{1}{2}(a_{11} - a_{22}) \cos 2\varphi \tag{33}$$

$$a_{pn} = \frac{1}{2}(a_{11} - a_{22}) \sin 2\varphi$$

or from (32),

$$a_{nn} = \frac{1}{2}(a_{11} + a_{22}) - \frac{1}{2}(a_{11} - a_{22}) \cos 2(\alpha - \beta) \tag{34}$$

$$a_{pn} = \frac{1}{2}(a_{11} - a_{22}) \sin 2(\beta - \alpha)$$

Equation 34 gives the values of a_{nn} and a_{pn} for a fixed value of β when α varies. The same procedure may be repeated for a second peri-

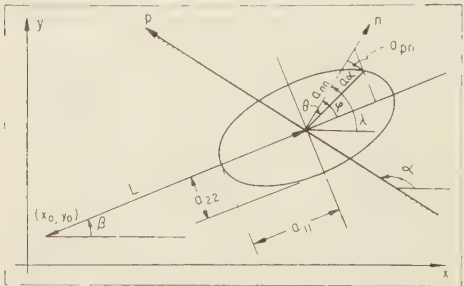


Fig. 7. The ellipse of variances for a uniform field of flow in the direction β .

with u_2 , β_2 , L_2 , $a_{11}^{(2)}$, etc. For the m th period one obtains

$$\begin{aligned} \sigma_n^{(m)} &= \frac{1}{2}(a_{11}^{(m)} + a_{22}^{(m)}) \\ &\quad - \frac{1}{2}(a_{11}^{(m)} - a_{22}^{(m)}) \cos 2(\alpha - \beta^{(m)}) \\ \sigma_n^{(m)} &= \frac{1}{2}(a_{11}^{(m)} - a_{22}^{(m)}) \sin 2(\beta^{(m)} - \alpha) \end{aligned} \quad (35)$$

here the superscript m denotes the m th period. If the same value of α is used for a series of n steps, i.e., the same plane p is displaced several times, the various values of $a_{nn}^{(m)}$ and $\sigma_n^{(m)}$ may be summed up to yield the corresponding variances of the final distribution. Hence

$$\begin{aligned} \sigma_n &= \frac{1}{2} \sum_{(m)} (a_{11}^{(m)} + a_{22}^{(m)}) \\ &\quad - \frac{1}{2} \sum_{(m)} (a_{11}^{(m)} - a_{22}^{(m)}) \cos 2(\alpha - \beta^{(m)}) \\ \sigma_n &= \frac{1}{2} \sum_{(m)} (a_{11}^{(m)} - a_{22}^{(m)}) \sin 2(\beta^{(m)} - \alpha) \end{aligned} \quad (36)$$

This operation is equivalent to the vector addition of all the vectors a_{α} associated with the same plane p in the various periods. Here it is done by adding the corresponding normal and tangential components.

When the angle α is the given angle of an initially abrupt interface, then $(a_{nn})^{\frac{1}{2}}$ is the standard deviation of the ultimate distribution perpendicular to the direction of the initial interface. When no such plane is given, the ultimate value of \bar{a}_{α} may be computed from \bar{a}_{nn} and \bar{a}_{pn} . The end points of \bar{a}_{α} will describe an ellipse as α varies between 0 and 2π . From this ellipse the ultimate distribution resulting from a point injection after an arbitrary series of movements can be determined. One can determine its major and minor axes and its orientation, and from these determine the actual ultimate tracer distribution. The above procedure is based on the fact that

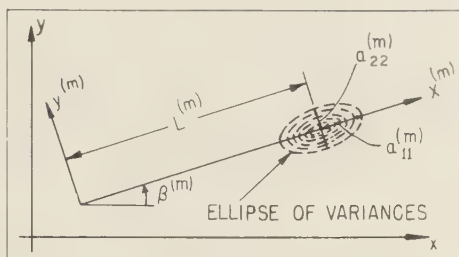


Fig. 8. The ellipse of variances for the m th period.

$$a_{ij} = \begin{bmatrix} a_{11}^{(m)} & 0 \\ 0 & a_{22}^{(m)} \end{bmatrix}; \quad \begin{aligned} a_{11}^{(m)} &= 2 D_{11} L^{(m)} \\ a_{22}^{(m)} &= 2 D_{11} L^{(m)} \end{aligned} \quad (37)$$

whose principal direction makes an angle $\beta^{(m)}$ with the positive x axis. Equation 37 can be written for each period in the sequence of r periods. The ultimate distribution is obtained by a summation of the r tensors. All tensors must be expressed in the same coordinate system before the summation. The transformation is performed by the rule

$$a_{st} = l_{si} l_{tj} a_{ij}; \quad i, j, s, t = 1, 2 \quad (38)$$

where a_{ij} denotes the tensor of (37), a_{st} denotes the same tensor in the x, y coordinates, and the summation convention is used. According to this convention, an index which is repeated twice indicates a summation over the entire range of that index (e.g., 1 and 2 in the two-dimensional case treated here). The quantities l_{si} and l_{tj} denote the cosines of the angles between the respective axes. These are $\cos \alpha$, $\sin \alpha$, $-\sin \alpha$, and $\cos \alpha$ for l_{11} , l_{12} , l_{21} , and l_{22} , respectively. Using the rule given by (38) and adding the resulting tensors, one obtains (39) for the distribution after r periods

$$\begin{bmatrix} \sum_{m=1}^r (a_{11}^{(m)} \cos^2 \beta^{(m)} + a_{22}^{(m)} \sin^2 \beta^{(m)}) & \sum_{m=1}^r (a_{11}^{(m)} - a_{22}^{(m)}) \sin \beta^{(m)} \cos \beta^{(m)} \\ \sum_{m=1}^r (a_{11}^{(m)} - a_{22}^{(m)}) \sin \beta^{(m)} \cos \beta^{(m)} & \sum_{m=1}^r (a_{11}^{(m)} \sin^2 \beta^{(m)} + a_{22}^{(m)} \cos^2 \beta^{(m)}) \end{bmatrix} \quad (39)$$

variances of normal distributions are additive.

Equation 36 can be obtained directly in view of the fact that a_{11} and a_{22} are the principal components of the tensor of variances. This is a symmetric tensor of rank 2 in Cartesian coordinates. Thus, the concentration distribution resulting from a point injection in the m th period (Figure 8) is described by the tensor

The principal values and the direction of the tensor (39) can be determined by making use of tensor algebra. If \bar{a}_{11} and \bar{a}_{22} are the principal values and β_0 is the angle between \bar{a}_{11} and the positive x axis, then

$$\tan 2\beta_0 = 2 \frac{\sum_{m=1}^r (a_{11}^{(m)} - a_{22}^{(m)}) \sin 2\beta^{(m)}}{\sum_{m=1}^r (a_{11}^{(m)} + a_{22}^{(m)}) \cos 2\beta^{(m)}}$$

Since, by definition,

$$a_{11}^{(m)} = 2 D_I L^{(m)}; \quad a_{22}^{(m)} = 2 D_{II} L^{(m)}$$

it follows that

$$\tan 2\beta_0 = \frac{\sum L^{(m)} \sin 2\beta^{(m)}}{\sum L^{(m)} \cos 2\beta^{(m)}} \quad (40)$$

and

$$\begin{aligned} \bar{a}_{11}, \bar{a}_{22} &= (D_I + D_{II}) \sum L^{(m)} \\ &\pm (D_I - D_{II}) \{ [\sum L^{(m)} \cos 2\beta^{(m)}]^2 \\ &+ [\sum L^{(m)} \sin 2\beta^{(m)}]^2 \}^{1/2} \end{aligned} \quad (41)$$

The final distribution is described by

$$\bar{a}_{ij} = \begin{bmatrix} \bar{a}_{11} & 0 \\ 0 & \bar{a}_{22} \end{bmatrix} \quad (42)$$

with \bar{a}_{11} in the direction defined by β_0 .

The expression under the root sign of (41) cannot, in general, be expressed as a function of the total path and/or the distance between the original point injection and its final position. This leads to the important conclusion that, in general, the final concentration distribution depends upon the exact path. Special cases are those in which all $\beta^{(m)}$'s are equal to each other or differ by π . The order of the steps themselves in the sum is immaterial.

Another conclusion from this discussion is that the ultimate distribution cannot be determined from the vertical and horizontal components of the velocity. This will be discussed further later, when it will be shown that displacements should be considered as second-rank tensors rather than as vectors.

The ultimate tracer distribution resulting from a point injection subjected to a sequence of uniform flows in various directions in a two-dimensional field of flow is given by

$$\begin{aligned} C(x, y) &= \frac{C_0}{2\pi\sigma_I\sigma_{II}} \\ &\exp \left\{ - \frac{[(y - \bar{y}) \sin \beta_0 + (x - \bar{x}) \cos \beta_0]^2}{2\sigma_I^2} \right. \\ &\quad \left. - \frac{[(y - \bar{y}) \cos \beta_0 - (x - \bar{x}) \sin \beta_0]^2}{2\sigma_{II}^2} \right\} \end{aligned} \quad (43)$$

where \bar{x} and \bar{y} are defined in Figure 9, β_0 is determined from (40), and $\sigma_I^2 = \bar{a}_{11}$; $\sigma_{II}^2 = \bar{a}_{22}$ are determined from (41). This result is obtained directly from the ellipse of the final tracer distribution, which is defined by (40) and (41). Equation 43 is the same as (12), obtained by direct integration.

When a certain region is initially occupied by a tracer fluid and the entire field is then subjected to a sequence of uniform movements, the final tracer distribution is obtained by integrating (43) over the entire region covered by the tracer fluid.

For dispersion in a three-dimensional isotropic field of flow, the procedure described above may be followed exactly. The dispersion tensor will be given by

$$a_{ij} = \begin{bmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{bmatrix} \quad (44)$$

with

$$a_{11} = 2 D_I L = \sigma_I^2, \quad (45)$$

$$a_{22} = a_{33} = 2 D_{II} L = \sigma_{II}^2 = \sigma_{III}^2$$

When tensor operations are performed, the indices cover the range 1, 2, 3.

Equations 44 and 45 indicate the dependence of the dispersion tensor a_{ij} upon the longitudinal and lateral constants of dispersion and upon the displacement.

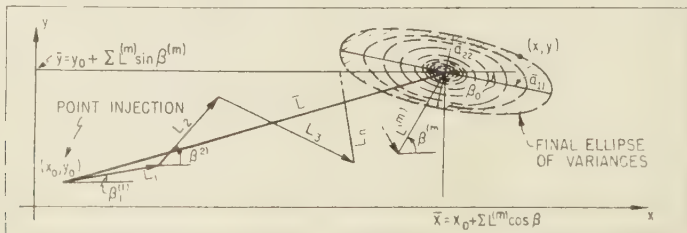


Fig. 9. The resulting ellipse of variances after a sequence of motions.

The constant of dispersion. Equation 45 shows that each component of the tensor of dispersion is twice the product of two factors, one describing the displacement and the other describing the dispersive property of the porous medium to disperse the tracer. It would therefore be natural to seek a general expression that will relate the tensor of dispersion a_{ij} to the dispersive property of the porous medium and to the displacement causing dispersion.

From the discussion presented in the previous section it follows that the components of the mean velocity vector ($u \cos \beta$ and $u \sin \beta$) or of the mean displacement vector ($L \cos \beta$ and $L \sin \beta$) play no role in the analysis. This is probably due to the nature of the porous medium, with its interconnected channels distributed at random in all directions in an isotropic medium, and to the fact that the true velocity in the channels causes the dispersion and not the mean velocity over the entire cross section and its components.

The dispersion in the direction of the mean displacement L (Fig. 10a) and perpendicular to it can be obtained by the longitudinal and lateral constants of dispersion, respectively, but no information is available on the dispersion resulting from the horizontal component or the vertical component of the flow in the x, y plane. If the displacement is referred to the x', y' coordinate system with x' being parallel to the direction of the mean flow, it can be defined by one component L (a vector in the direction of the flow) and another component of magnitude zero in the direction perpendicular to the flow.

Therefore, although the vector of mean displacement L can be separated into its components $L \cos \beta$ and $L \sin \beta$, to obtain the dispersion itself each component has to be further resolved to yield displacement vectors in the directions of the mean flow and perpendicular to it. The vector \vec{AB} in Figure 10b is equivalent to the sum

of the displacements \vec{AC} and \vec{CB} which are the components of the displacement \vec{AB} in the x and y directions, respectively. Each of these two components is in turn equivalent to the vector sum of two displacements of which one is in the direction of the mean flow and the other is perpendicular to it. The vector-sums can be written:

$$\begin{aligned}\vec{AB} &= \vec{AC} + \vec{CB} \\ \vec{AC} &= \vec{AD} + \vec{DC} \\ &= L \cos^2 \beta + L \sin \beta \cos \beta \quad (46) \\ \vec{CB} &= \vec{CD} + \vec{DB} \\ &= L \sin^2 \beta + L \sin \beta \cos \beta\end{aligned}$$

Vectors \vec{DC} and \vec{CD} are equal in magnitude and opposite in direction, so that they cancel each other, leading to the result

$$\begin{aligned}\vec{AB} &= \vec{AD} + \vec{DB} \\ &= L \cos^2 \beta + L \sin^2 \beta = L \quad (47)\end{aligned}$$

The four components $L \cos^2 \beta$, $L \sin \beta \cos \beta$, $L \cos \beta \sin \beta$, and $L \sin^2 \beta$ thus define the displacement in the x, y coordinates by means of displacements in the direction of the flow and perpendicular to it. The same displacement in the x', y' coordinates has the components $L, 0, 0, 0$. Hence the displacement as a physical quantity transforms from one system of coordinates to another according to the rule given by (38). This statement means that the displacement in the porous medium should be regarded as a second-rank tensor.

If the second-rank tensor of displacement is denoted by L_{kl} , then, when the mean displacement $L = ut$ is in the x' direction, the displacement in the x', y' coordinates is expressed by

$$L_{kl} = \begin{bmatrix} L & 0 \\ 0 & 0 \end{bmatrix} \quad k, l = 1, 2 \quad (48)$$

In general, when the mean flow makes an angle β with the x axis, the displacement is given by

$$L_{kl} = \begin{bmatrix} L \cos^2 \beta & L \sin \beta \cos \beta \\ L \sin \beta \cos \beta & L \sin^2 \beta \end{bmatrix} \quad (49)$$

Equation 49 is obtained by rotating the coordinates of (48) according to the rule given by 38. When a three-dimensional field of flow is con-

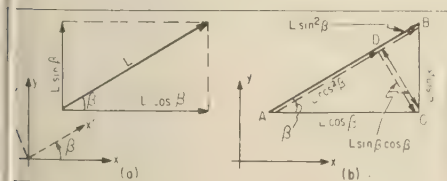


Fig. 10. The resolution of the displacement L .

sidered, the indices k, l take on the values 1, 2, 3.

The fact that in this case the displacement should be considered a tensor rather than a vector can be attributed to the channel-system character of porous media. Because of these channels, a hydraulic gradient in a certain direction will cause microscopic flows in the channels in all directions. The dispersion phenomena are actually caused by these microscopic flows. The lateral flow caused by the vector grad φ , with φ being the potential, suggests an analogy with the theory of elasticity, in which a force in a certain direction also causes deformation in the lateral direction. The strain, which is analogous to the displacement here, is also a second-rank tensor.

The relationship between a_{ij} and L_{kl} could be written in the form

$$a_{ij} = 2 D_{II} L \delta_{ij} + 2(D_I - D_{II})L_{kl} \quad (50)$$

where δ_{ij} is the Kronecker symbol.

Another way of expressing the relationship between the dispersion described by a_{ij} and the displacement defined by L_{kl} is when the constant coefficient is defined as a tensor of the fourth rank. If this constant of dispersion is denoted by D_{ijkl} , then in the general case of a mean flow in the β direction in an isotropic two-dimensional porous medium the relationship between the variances of the dispersion a_{ij} , the displacement L_{kl} , and the constant of dispersion D_{ijkl} is given by

$$a_{ij} = 2 D_{ijkl} L_{kl}; \quad i, j, k, l = 1, 2 \quad (51)$$

For the three-dimensional case, $i, j, k, l = 1, 2, 3$.

In general, a fourth-rank tensor in three dimensions contains 81 components (or 16 in two dimensions) of which all but 21 (or 8 in two dimensions) are equal to zero. The components which differ from zero are those with four equal indices or two pairs of equal indices. For isotropic

D_I, D_{II} = longitudinal and lateral constant of dispersion of the (isotropic) porous medium (dimension, cm).

a_{ij} = the variance which is a second-rank tensor (dimension, cm²).

D_{ijkl} = the constant of dispersion of the porous medium, or dispersivity of the porous medium. This is a fourth-rank tensor (dimensions, cm).

The components of D_{ijkl} , in general, depend on the chosen frame of reference, inasmuch as the variance a_{ij} and the displacement vary with the coordinate system. For an isotropic medium, however, the components of D_{ijkl} remain invariant under any transformation of the Cartesian coordinates.

A compact form in which to write a fourth-rank tensor in two dimensions would be the following:

$$D_{ijkl} = \begin{bmatrix} A_{1111} & A_{1122} & A_{1112} & A_{1121} \\ A_{2211} & A_{2222} & A_{2212} & A_{2221} \\ A_{1211} & A_{1222} & A_{1212} & A_{1221} \\ A_{2111} & A_{2122} & A_{2112} & A_{2121} \end{bmatrix} \quad (52)$$

The constant of dispersion is then given by

$$D_{ijkl} = \begin{bmatrix} D_I & D_{II} & 0 & 0 \\ D_{II} & D_I & 0 & 0 \\ 0 & 0 & (D_I - D_{II})/2 & (D_I - D_{II})/2 \\ 0 & 0 & (D_I - D_{II})/2 & (D_I - D_{II})/2 \end{bmatrix} \quad (53)$$

Thus, the dispersion (54) from a point injector in a uniform field of flow making an angle β with the x axis is obtained directly from (51) along with (49) and (53):

$$a_{ij} = 2 \begin{bmatrix} L D_I \cos^2 \beta + L D_{II} \sin^2 \beta & L(D_I - D_{II}) \sin \beta \cos \beta \\ L(D_I - D_{II}) \sin \beta \cos \beta & L D_I \sin^2 \beta + L D_{II} \cos^2 \beta \end{bmatrix} \quad (54)$$

porous media, all these components depend on two essential constants which are the longitudinal and lateral constants of dispersion of the porous medium.

The following summary of nomenclature is suggested here to avoid confusion:

Equation 54 is the same as (39), obtained previously in a different way. The major and minor axes of the ellipse described by (54) are $2D_I L$ and $2D_{II} L$, respectively, with $2D_I L$ in the direction of the mean flow. The same result could be obtained by using (50).

The combined effect of a sequence of steps is given by

$$\bar{a}_{ij} = \sum_{(m)} a_{ij} = 2 \sum_{(m)} (D_{ijk} L_{kl}) \quad (55)$$

For an isotropic medium, D_{ijk} is the same for all steps; hence

$$\bar{a}_{ij} = 2 D_{ijk} L_{kl}; \quad L_{kl} = \sum_{(m)} L_{kl}^{(m)} \quad (56)$$

$$\left[\frac{\sum L^{(m)} \cos^2 \beta^{(m)}}{\sum L^{(m)} \sin 2\beta^{(m)}} \frac{1}{2} \frac{\sum L^{(m)} \sin 2\beta^{(m)}}{\sum L^{(m)} \sin^2 \beta^{(m)}} \right] \quad (57)$$

An equivalent path that will yield the same tracer distribution may be determined. The equivalent displacement is given by

$$L_{kl}^0 = \begin{bmatrix} L_{11}^0 & 0 \\ 0 & L_{22}^0 \end{bmatrix} \quad (58)$$

oriented in the direction β_0 with respect to the positive x axis, with

$$\tan 2\beta_0 = \frac{\sum L^{(m)} \sin 2\beta^{(m)}}{\sum L^{(m)} \cos 2\beta^{(m)}} \quad (59)$$

and

$$L_{11}^0, L_{22}^0 = \frac{1}{2} \sum L^{(m)} \pm \frac{1}{2} \left[\left(\sum L^{(m)} \cos 2\beta^{(m)} \right)^2 + \left(\sum L^{(m)} \sin 2\beta^{(m)} \right)^2 \right]^{1/2} \quad (60)$$

The result emphasizes the tensorial nature of the summation. When a final distribution is given, L_{22} will be equal to zero if all β 's are equal to a constant (or a constant $\pm\pi$), or if the ratio of the major to the minor axis of the ellipse equals the ratio D_I/D_{II} .

The analysis could be carried out in a similar manner for the three-dimensional case.

The aim stated in the beginning of this work was thus achieved. The property of an isotropic porous medium to disperse the tracer fluid was separated from the flow and from the resulting dispersion. The property which is defined by the constant of dispersion D_{ijk} depends only upon the characteristics of the porous medium and the

geometry of its pore-channel system. These characteristics are expressed by the longitudinal and lateral constants of dispersion of the porous media. A second paper in which experiments in dispersion confirm some of the statements made in this paper is now being prepared for publication.

The problem of dispersion in a nonisotropic material still remains unsolved. One should distinguish between various kinds of nonisotropy and repeat the statistical analysis suggested by de Josselin de Jong [1958] and Saffman [1959] with different frequency functions for the spatial distribution of channels in each case. This might yield a different relationship between the dispersive property of the porous media, the displacement, and the resulting dispersion.

Acknowledgment. I wish to thank Dr. G. de Josselin de Jong, Delft Technical University, for his helpful suggestions and review of this work at all stages, Professor D. K. Todd, University of California, Berkeley, for reviewing this work, and Dr. A. Gutwirth, Israel Institute of Technology, for his mathematical advice.

REFERENCES

- Bear, J., The transition zone between fresh and salt waters in coastal aquifers, Ph.D. thesis, Civil Eng., Univ. Calif., Berkeley, 139 pp., 1960.
- de Josselin de Jong, G., Longitudinal and transverse diffusion in granular deposits, *Trans. Am. Geophys. Union.*, 39, 67-74, 1958.
- Feller, W., *An Introduction to Probability Theory and Its Applications*, John Wiley & Sons, New York, 461 pp., 1957.
- Lau, L. K., Dispersion of water tracer in radial laminar flow through porous media, Ph.D. thesis, Univ. Calif., Berkeley, 1959.
- Lau, L. K., W. J. Kaufman, and D. K. Todd, Dispersion of a water tracer in radial laminar flow through porous media, *Inst. Eng. Research Rept.* 93-5, Univ. Calif., Berkeley, 78 pp., 1959.
- Rifai, E. M. N., W. J. Kaufman, and D. K. Todd, Dispersion phenomena in laminar flow through porous media, *Progr. Rept. 2*, Sanitary Eng. Research Lab., Univ. Calif., Berkeley, 156 pp., 1956.
- Saffman, P. G., A theory of dispersion in porous medium, *J. Fluid Mech.*, 6, 321-349, 1959.
- Scheidegger, A. E., Statistical approach to miscible displacement in porous media, *Can. Mining Met. Bull.*, 52, 26-30, 1959.

(Manuscript received July 12, 1960; revised December 23, 1960.)



Capillary Pressure and Surface Discontinuity in Porous Media

RAO CHANNAPRAGADA AND WALTER ROSE

*University of Illinois
Urbana, Illinois*

Abstract. A derivation of the capillary pressure term for viscous fluids at a fluid-fluid interface is given in the general form. *Hall's* [1956] paper is discussed in part, and it is emphasized that the surface of discontinuity between the fluids cannot be ignored in describing the dynamics of multiphase fluid flow through a porous medium

Notation.

F_t	Tangential force
n_i	Normal unit vector
n_k	Tangential unit vector
p_c	Capillary pressure force
p_1, p_2	Ordinary pressure
R_1, R_2	Radii of curvature
S	Body force per unit mass
v	Velocity vector of the fluid
σ_{ik}	Stress tensor
σ_{ik}'	Viscous stress
γ	Surface tension
μ_i	Fluid viscosity
ρ_i	Fluid density
∇	Gradient (grad)
θ	Surface concentration

The equations of motion of an element of fluid in a two-phase (D_1 and D_2) flow through a porous medium (Fig. 1) are given, in a microscopic sense, by the well-known Stokes-Navier equations [Hubbert, 1956]:

$$\rho_1 \frac{dv_1'}{dt} = S\rho_1 - \nabla p_1' + \mu_1 \nabla^2 v_1' \quad (1)$$

and

$$\rho_2 \frac{dv_2'}{dt} = S\rho_2 - \nabla p_2' + \mu_2 \nabla^2 v_2' \quad (2)$$

The boundary conditions that must be met at the interface between the two phases are that velocities v_1' and v_2' are equal and that, if the interface is stationary, the forces exerted on the interface are equal and opposite. That is,

$$[v_1' = v_2']_{\text{interface}} \quad (3)$$

and

$$n_{1,k}\sigma_{1,i,k} + n_{2,k}\sigma_{2,i,k} = 0 \quad (4)$$

The normal vectors n_1 and n_2 are in opposite direction, i.e.,

$$n_{1,i} = -n_{2,i} = n_i \quad (5)$$

so that (4) can be written as

$$n_i\sigma_{1,i,k} = n_i\sigma_{2,i,k} \quad (6)$$

The capillary pressure at the interface is defined as

$$\begin{aligned} P_c &= (p_1' - p_2')n_i \\ &= \gamma(1/R_1 + 1/R_2)n_i - \sigma_{ik}'n_k \end{aligned} \quad (7)$$

In the case of ideal fluids the viscous stresses, $\sigma_{ik}'n_k = (\sigma_{1,i,k} - \sigma_{2,i,k})n_k$, are zero, and (7) reduces to the classical differential equation

$$P_c = \gamma(1/R_1 + 1/R_2)n_i \quad (8)$$

The surface tension γ , in general, is not a constant over the surface of separation but is a function of the temperature and of surface concentration θ of the adsorbed film. In addition to the normal force (which is zero for a plane interface), there is a tangential force ($F_t = \text{grad } \gamma$)

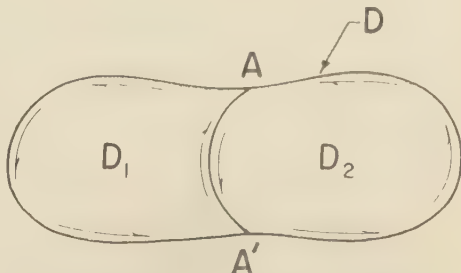


Fig. 1. Element of a fluid in a two-phase flow through a porous medium.

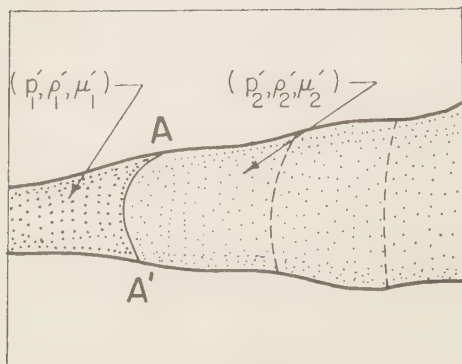


Fig. 2. Motion of an interface in two-phase flow.

per unit area of the surface of separation [Rose and Channapragada, 1960].

The gradient of surface tension is taken as positive, unlike the pressure, since the surface tension forces tend to reduce the surface area whereas the pressure forces tend to increase the volume.

In a pore of an arbitrary shape (Fig. 2), as the interface advances, the shape of the interfacial surface changes with the motion. The film at the interface is stretched or compressed depending on the pore geometry, and thereby the surface concentration of the adsorbed substance is changed. Then the tangential force F_t , in the presence of an adsorbed film at the interface, is given by

$$F_t = \frac{\delta\gamma}{\delta\theta} \text{grad } \theta \quad (9)$$

Adding to (7) the term defined in (9), one obtains for capillary pressure

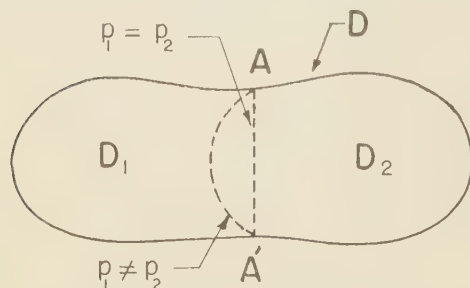


Fig. 3. Liquid-gas interface replaced by liquid-liquid interface having the same pressures on both sides of the interface.

$$P_c = \gamma \left(\frac{1}{R_1} + \frac{1}{R_2} \right) n_i + (\sigma'_{2,ik} - \sigma'_{1,ik}) n_k + \frac{\delta\gamma}{\delta\theta} \text{grad } \theta \quad (10)$$

Hall [1956], in his derivation of the Darcy law hypothetically replaced the liquid-gas interface by a liquid-liquid interface having the same pressures on both sides of the interface (Fig. 3). He further claimed that 'the force due to surface tension on AA' is replaced by liquid pressure without any change on the state of stress on the fluid within the region previously bounded by the meniscus. The force system of Figure 1 may thus be replaced by the equivalent liquid pressure system of Figure 3, without changing the total force per unit volume due to these causes.' Hall's analysis is equivalent to stating that the sum of the integrals around D_1 and D_2 in the plane is equal to the contour integral around D , or

$$\oint_{D_1} f_1 ds + \oint_{D_2} f_2 ds = \oint_D f ds \quad (11)$$

The equality of the integrals would be valid provided that there is no essential singularity at the interface. In the case of real functions all it means is that the derivatives fail to exist because of the discontinuity of the functions at the interface. Hence one cannot represent the internal forces by the forces acting on the boundary L when there is a surface discontinuity inside of D .

It is the opinion of the present authors that by such replacement as was mentioned by Hall a certain amount of the physical significance of the problem is lost.

In conclusion, the interfacial discontinuity cannot be artificially removed, as was done by Hall, since the singularity at the interface is nonremovable. Thus, in order to describe the motion of a particle one has to apply equation 1 and 2 separately to each of the fluid phase with the above-stated boundary conditions at the interface.

Further, during the fluid motion, the interface is not a constant curvature surface and can be expressed in general form as a function of the following variables:

$$(1/R_1 + 1/R_2) = G(p, \sigma, \gamma, \theta, \delta\gamma/\delta x_i) \quad (12)$$

Acknowledgment. This work was supported by funds received from the Chemical Warfare Laboratories CML517.

REFERENCES

all, W. A., Analytical derivation of the Darcy equation, *Trans. Am. Geophys. Union*, 37, 185-188, 1956.

ubbert, M. King, Darcy law and the field equa-

tions of the flow of underground fluids, *AIME J. Petrol. Technol.*, p. 222, October 1956.

Rose, Walter, and Rao Channapragada, Microscopic aspects of capillary imbibition, *AIME Paper 1549-G*, presented at the Society of Petroleum Engineers, Denver, Colorado, October 1960.

(Manuscript received August 3, 1960; revised January 4, 1961.)

Estimating Ground-Water Recharge from Stream Hydrographs¹

P. MEYBOOM

*Groundwater Division, Research Council of Alberta
Edmonton, Alberta, Canada*

Abstract. It is possible to separate graphically the base flow on stream hydrographs by plotting the logarithm of the discharge against time. Total potential ground-water discharge (Q_{tp}) at the beginning of any given base-flow recession is

$$Q_{tp} = K_1 K_2 / 2.3$$

where K_1 = ground-water discharge at the beginning of the base-flow recession (t_0), in acre feet per day; K_2 = time increment corresponding to one log cycle change in Q , in days.

The difference between the actual amount of ground-water discharge at the end of the base-flow recession and the total potential ground-water discharge at the beginning of the same recession is called 'remaining potential ground-water discharge.' The difference between the total potential ground-water discharge at the beginning of any given base-flow recession and the remaining potential ground-water discharge at the end of the preceding base-flow recession is a measure of ground-water recharge. The method is illustrated by the computation of the ground-water balance of the Calgary area (Alberta, Canada).

INTRODUCTION

During the course of a hydrogeological survey of the Calgary area, Alberta, Canada, a method of hydrograph analysis was devised to differentiate among the various components of stream runoff. The twofold purpose of the study was to obtain quantitative information concerning the basic hydrologic equation: ground-water recharge = ground-water discharge \pm change in ground-water storage, and to determine to what extent the various types of aquifers in this area are affected by natural recharge.

Calgary (51°N, 114°W) is situated at the confluence of the Bow and Elbow rivers, east of the Rocky Mountains (Fig. 1). Geologically the area lies near the western margin of the Western Canada Sedimentary Basin and east of the Foothills Belt. Nonmarine sandstones and shales of the Paleocene Paskapoo formation constitute the bedrock of the Calgary area, and are, in turn, covered by a mantle of glacial drift ranging in thickness from 1 to 100 feet. Ground water in the area is obtained from three different sources: thick alluvial gravel deposits along the rivers, aquifers in the glacial drift, and the bedrock. Natural discharge from each of these sources contributes to streamflow, and the nature and

magnitude of these contributions has been determined by stream hydrograph analyses. Before discussing the results obtained in calculating the ground-water balance in the area, the method of analysis will be explained in detail.

COMPONENTS OF STREAM RUNOFF

It has long been recognized that stream runoff consists of three components [Barnes, 1939; Meyer, 1940] which are related to the three different ways in which rain water reaches streams. These components are (a) *direct runoff* which flows over the land surface, (b) *interflow* (or secondary base flow) which consists of water that flows part of the way underground but does not become part of the main ground-water body, and (c) *base flow* which is natural ground-water discharge. Direct runoff occurs only during and immediately after periods of rain; interflow generally continues somewhat longer, but during extended periods of dry weather stream flow is maintained solely by ground-water discharge [Barnes, 1939; Linsley, Kohler, and Paulhus, 1949; Natermann, 1950] and in some rivers by melted snow [Laycock, 1957].

Base flow, also called ground-water recession or ground-water discharge, represents withdrawal of ground water from aquifer storage after ground-water recharge has ceased. As ground-water discharge alone depends mainly

¹ Research Council of Alberta Contribution 40.

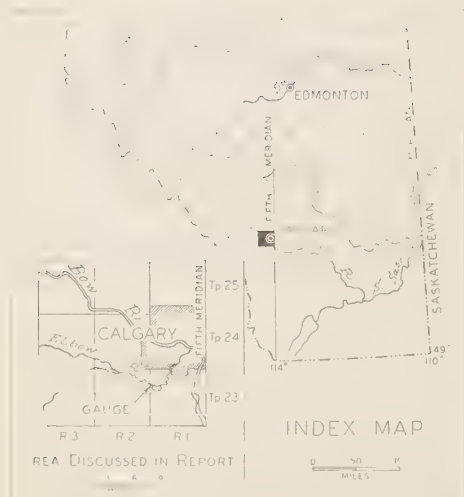


Fig. 1. Map of southern Alberta, showing location and general features of the Calgary area.

on the physical and geological properties of the drainage basin, its magnitude gives important information as to the permeability and storage capacity of the aquifers in the basin.

Runoff is generally shown graphically in stream hydrographs, and several graphic techniques have been developed to separate the various flow components [Linsley, Kohler, and Paulhus, 1949]. Each technique has certain advantages depending on the type of investigation for which it is to be used, and depending also on the geology of the drainage basin.

Barnes [1939] found that the recession of each of the hydrograph components can be expressed empirically by the general regression equation

$$Q_t = Q_0 K_r^t \quad (1)$$

where

Q_0 = discharge at any given time, in acre feet per day.

Q_t = discharge t time units after Q_0 , in days.

K_r = daily recession factor.

t = time interval, in days.

If the logarithm of the discharge is plotted against time, a linear relationship between the two variables may be obtained. The best straight line is generally for the ground-water recession.

To calculate the amount of natural ground-water discharge that can be considered typical

of the Calgary area, the hydrograph of the Elbow River above Glenmore Dam (Fig. 1) was plotted semilogarithmically (Figs. 5 and 6) for 9 successive years. The drainage area above the station comprises 460 square miles, consisting of the outer front ranges and foothills of the Rocky Mountains. The upper part of the basin has relatively large areas of bare rock, and runoff is rapid. Stream flow is maintained by melting snow in May, June, and July as well as by rainfall, which reaches a maximum in June. The lower part of the basin receives some additional flow from above-normal spring rains in April and May [Laycock, 1957]. The average daily ground-water recession factor K_r for this station in the Elbow River was calculated to be 0.9883, with a minimum value of 0.9785 during 1950-1951 and a maximum value of 0.9915 during 1955-1956.

COMPONENTS OF BASEFLOW

Before explaining in more detail the method by which the Elbow River hydrograph was analyzed, it is necessary to recognize the various sources of ground water contributing to base flow. As can be seen from Figure 2, ground water from three different origins is discharged into the river:

Contact springs. The surficial deposits covering the highlands along the Bow and Elbow rivers are all of glacial origin and consist of an unsorted mixture of boulders, sand, silt, and clay. These deposits are fairly permeable and have a moderate infiltration capacity. Rain water in filtrating into the drift cover will percolate downward until it reaches the underlying bed rock, which generally consists of interbedded

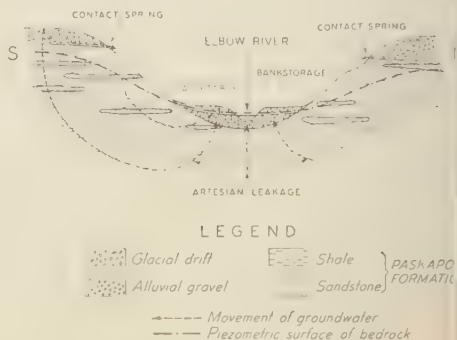


Fig. 2. Diagrammatic cross section of Elbow River valley, showing origin of baseflow components.

andstone and shale. The less permeable bedrock reverts part of the water from moving farther downward, thus deflecting it to the surface. Springs issuing in this fashion are called 'contact springs' [Meinzer, 1923]. All along the Bow and Elbow valleys the contact between drift and bedrock is marked by the presence of such springs, which generally yield less than 10 gallons per minute.

Artesian leakage. Water penetrating into bedrock becomes part of an artesian system that, despite apparent inhomogeneity of the bedrock, acts as a single hydrologic unit. This type of artesian system is common in western Alberta [Foster and Farvolden, 1958; W. A. Meneley, 1960, personal communication], and it is thought to be comparable to that described by [Hubbert 1940, p. 914, Fig. 45, p. 930]. From the shape of the piezometric surface (Fig. 8) it is obvious that the river is an area of natural ground-water discharge. This discharge process is called effluent artesian leakage.'

Bank storage. The third and probably the most important contribution of ground water to stream flow is generally referred to as 'bank storage' [Todd, 1955, 1959]. During flood periods the river becomes influent, thus forcing water to enter the gravel deposits along the stream. It has been shown [Todd, 1955] that large changes in magnitude and direction of ground-water movement are brought about during periods of river inefficiency. When the river level drops below the adjoining water table, the stream becomes effluent again, and water held in bank storage is released gradually.

The amount of possible recharge from rain falling on the flood plain is considered negligible, for, according to [Laycock 1957, Fig. 10; 1960, Fig. 3], the normal water deficiency in this area occurring between April and October is 3.7 inches. Occasionally a small water surplus may be produced by above-normal spring rains, but, as these occur during March or April when the soil is still frozen, their effect is small or altogether lost.

A GENERAL QUALITATIVE REVIEW OF FLUCTUATIONS IN BASE FLOW

Like direct runoff and interflow, base flow is susceptible to some fluctuations, but the magnitude of these fluctuations is considerably less than the variations in the other flow components.

The cause and character of the fluctuations of the various base-flow components will first be discussed qualitatively.

It follows from their nature that contact springs are somehow related to precipitation, but this relationship is generally somewhat subdued and is only slightly noticeable seasonally. As the discharge from contact springs is governed by the hydraulic gradient in the drift, summer discharge will be less than that during the other periods of the year owing to the higher evapotranspiration losses during the summer season. Furthermore, it was learned from inhabitants of the Springbank area (Fig. 2; Tp. 24, R. 2) that contact springs were particularly strong before 1920; in some cases their yield was even sufficient for irrigation. However, during the period 1920 to 1930 the flow decreased, and many springs dried up entirely between 1930 and 1935. A considerable time after 1935 some of the springs began to flow anew, but it was not until 1957 that all springs flowed as they had before the first drought. At the present time [1960] several springs that were flowing a few years ago have dried up again.

Figure 3, showing the cumulative departure from the mean total annual precipitation in Calgary over the past 60 years, clearly indicates these long-term fluctuations in precipitation. Each rising limb of the graph indicates a succession of years with above-average precipitation, whereas the descending limbs indicate years with below-normal values. [Wenzel 1936] demonstrated that long-term fluctuations in ground-water levels closely coincide with this type of long-term climatologic variation; therefore, the history of the contact springs in the Calgary area can also be read from this diagram (Fig. 3). It is obvious from this record that the flow of contact springs is influenced mainly by long-term variations in precipitation and only to a lesser degree by seasonal fluctuations in rainfall.

Artesian leakage, which is the second contribution to base flow, is governed by the transmissibility of the bedrock, the gradient of the piezometric surface, and the length of the discharge area. The only factor that may vary with time is the shape of the piezometric surface. It has been shown by [Jones 1960] and P. A. Domenico (1960, personal communication) that bedrock aquifers in Alberta are recharged by local precipitation. The greater the precipitation,

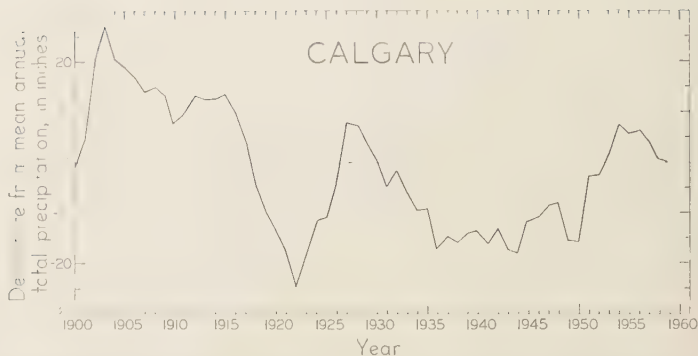


Fig. 3. Graph showing cumulative departure from mean annual total precipitation in Calgary, 1900-1959.

the greater the head difference between the drift aquifers and the bedrock and, consequently, the greater the recharge to bedrock. The shape of the piezometric surface will therefore also be influenced by long-term as well as seasonal fluctuations in precipitation. Artesian leakage is generally fairly diffuse, although artesian springs do exist.

The variations in the third component of base flow are of somewhat different character, owing to the full hydrologic connection between the stream and the gravels along the stream channel and to the greater permeability of these gravels. During the major part of the year the stream receives ground water discharged from the gravel banks, but during the flood periods, when the river level rises more rapidly than the adjacent ground-water table, the river becomes influent. This condition lasts until the river level again

drops below the adjoining water table. The amount of water released from bank storage during flood periods is uncertain, but it has been shown by Todd [1955] that the time required to re-establish ground-water discharge after a period of river effluency is extremely short. The peak runoff is known to pass through the channel in a few days, and this is immediately followed by an increase in ground-water flow from bank storage. Therefore, although ground water being released from bank storage in the vicinity of the gaging station is virtually nil at the time the peak runoff is recorded at the station, an increase in ground-water discharge from bank storage is already in effect upstream, resulting in a net increase in base flow at the gaging station as well. It is therefore likely that every major flood or flood period has an immediate effect upon the amount of base flow, and ground-water discharge

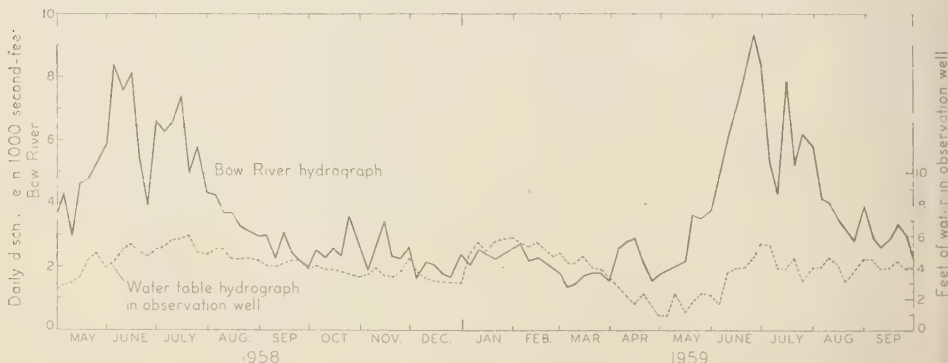


Fig. 4. Hydrographs showing total stream discharge of Bow River and fluctuations of the adjacent water table.

bank storage consequently varies appreciably with variations in precipitation.

Figure 4 shows how the fluctuations of the water table adjacent to the Bow River just below the confluence with the Elbow River are intimately linked with variations in stream discharge. It also shows that during this particular year, ground water was depleted from July 1958 until December 1958, was replaced slightly in January 1959, and was again depleted from January 1959 until the end of April 1959. Coincident with the very first spring flood the water table rises, indicating an immediate cessation of the baseflow recession, as has been suggested in the foregoing paragraph.

METHOD OF ESTIMATING GROUND-WATER DISCHARGE

It has been explained earlier that the part of the stream hydrograph that represents the ground-water recession becomes a straight line if the hydrograph is plotted semilogarithmically.

Figures 5 and 6 show hydrographs for 9 successive years for the Elbow River in which base flow has been separated from the other runoff components.

In estimating the amount of ground-water discharge the following assumptions are made:

(a) The straight lines connecting successive points of minimum stream discharge in Figures 5 and 6 are considered to approach true baseflow conditions. Slight upward deviations from this line are not caused by an increment in base flow but rather by slight increases in direct runoff produced by melting snow or occasional rains.

(b) The annual fluctuations in base flow are for the greater part related to changes in bank storage, compared to which the annual fluctuations in the contribution from contact springs and artesian leakage are only of lesser significance.

(c) Rainfall on the flood plain is considered a negligible source of recharge to bank storage, as the basin is located in a semiarid area which has

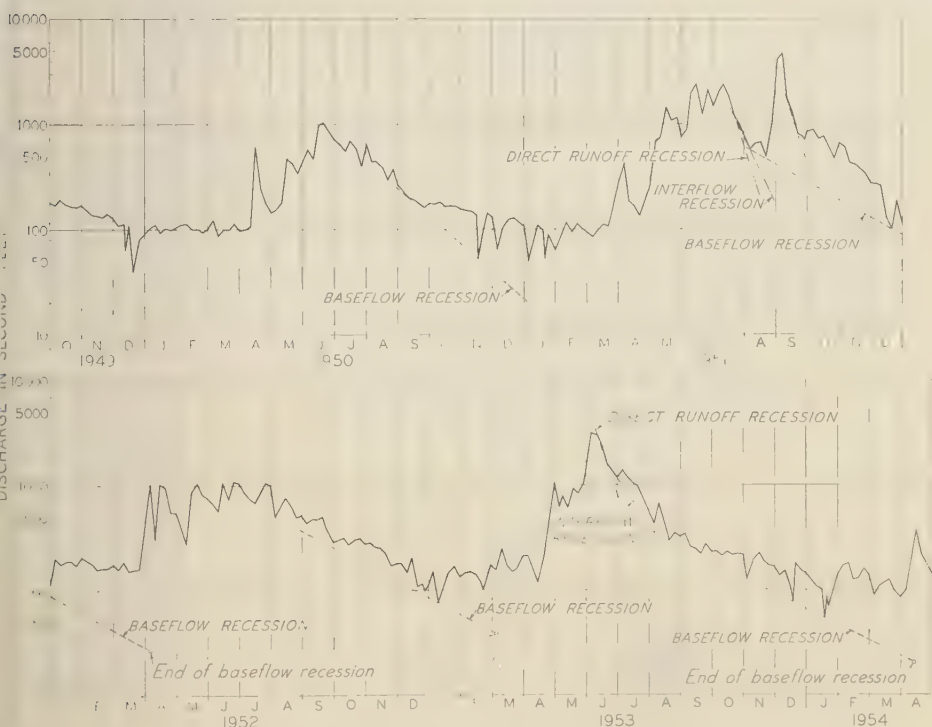


Fig. 5. Stream hydrograph of Elbow River, 1949-1954.



Fig. 6. Stream hydrograph of Elbow River, 1954-1959.

an average water deficiency of approximately 4 inches.

(d) The aggregate effect of the summer runoff and rainfall on all components of ground-water discharge is most noticeable toward the end of the summer, generally sometime in August. After this time the river level drops below the adjoining water table, and ground-water recession begins and continues until the next spring.

(e) Ground-water recession ceases as soon as one of the three base-flow components becomes affected by recharge. The first component to be affected is bank storage; consequently, true base-flow recession ceases with the first spring flood, generally sometime in April. It is realized that after this time ground-water contribution from contact springs and artesian leakage continues; however, as this continuation represents the very last stages of the individual recessions of these components, the error introduced in this manner is of the order of 0.1 inch, which, considering the nature of the analysis, seems to be acceptable.

The total volume of base flow still held in storage, corresponding to a given point on the

ground-water recession curve, can be calculated [Butler, 1957, eq. 10-7, p. 217]. The total volume of base flow that would be discharged during an entire ground-water recession, if complete depletion were to take place uninterruptedly, can be computed by integrating Butler's equation from time t_0 to infinity. This volume will be called 'total potential ground-water discharge' (Q_{tp}) and is given by

$$Q_{tp} = \int_{t_0}^{\infty} Q dt \quad (2)$$

$$= \left[\frac{-K_1 K_2 / 2.3}{10^{u/K_2}} \right] - \left[\frac{-K_1 K_2 / 2.3}{10^{0/K_2}} \right]$$

in which the first member becomes extremely small as t_2 approaches infinity, giving

$$Q_{tp} = - \left[\frac{-K_1 K_2 / 2.3}{10^{0/K_2}} \right]$$

in which $10^{0/K_2}$ equals unity. Thus the equation for total potential ground-water discharge can be written as

$$Q_{tp} = K_1 K_2 / 2.3 \quad (3)$$

in which K_1 and K_2 are as defined by *Butler* [1957, p. 216]. It appears that (3) is not a general equation but might be considered a specific solution of *Butler's* equation 10-9, p. 217.

The beginning of the ground-water recession, t_0 , is generally easy to determine, either by intersection with the interflow recession (Fig. 5, 1951, 1953; Fig. 6, 1955) or simply by taking the first minimum value that lies on the base-flow recession curve.

All base-flow recessions of the Elbow River appear to be virtually the same, with the exception of that of 1957-1958 (Fig. 6). However, meteorological data for this winter show that the average temperature during December 1957 was 28°F, which is 10° above the average December temperature. During this mild winter a considerable increase in direct runoff, probably resulting from melting snow, completely masked the base-flow recession. However, knowing the average daily recession factor K_r , it is possible to construct the base-flow recession curve for the winter of 1957-1958, which is shown by a dotted line in Figure 6. Closer examination of the other base-flow recessions reveals that the average recession graph slopes 6 months per log cycle, corresponding to an average daily recession

factor of 0.9883. Deviations from this slope indicate that ideal base-flow conditions are not reached during each recession period but that they are sometimes masked by continuing snow melt or excessive freezing, a condition which indeed introduces some uncertainty. However, these deviations appear to have but little influence, for the successive recharge values are sufficiently consistent to make the results acceptable as an over-all estimate.

The total potential ground-water discharge has been calculated for the beginning of each base-flow recession, and the results of these calculations are presented in Table 1, expressed both in acre feet and in depth in inches on the drainage area.

It follows from the nature of ground-water movement that the bulk of the total potential ground-water flow will be discharged during the first months of the base-flow recession. The total volume of ground water actually discharged during any given ground-water recession has been calculated from *Butler's* equation 10-6. This volume will be called 'actual ground-water discharge.' The volumes of actual ground-water discharges during the eight ground-water recessions are also given in Table 1. It has been

TABLE 1. Elbow River Hydrograph Analysis, 1950-1958

Duration of Ground- Water Recession	Total Potential Ground-Water Discharge		Actual Ground- Water Discharge		Per- centage of Total Flow	Remaining Po- tential Ground- Water Discharge		Total Recharge, inches	Change in Storage, inches
	acre-feet	inches	acre-feet	inches		acre-feet	inches		
Aug. 20, 1950	31,339	1.27							
Mar. 30, 1951			31,114	1.26	53	225	0.01	1951: 3.81	
Aug. 5, 1951	93,848	3.82							+2.55
Apr. 5, 1952			89,511	3.64	30	4337	0.18	1952: 3.13	
Aug. 5, 1952	81,261	3.31							-0.51
Mar. 10, 1953			75,403	3.07	58	5858	0.24	1953: 2.22	
Aug. 7, 1953	60,358	2.46							-0.85
Apr. 15, 1954			56,709	2.31	52	3649	0.15	1954: 2.62	
July 30, 1954	67,978	2.77							+0.31
Apr. 10, 1955			62,514	2.54	37	5464	0.23	1955: 1.28	
Aug. 20, 1955	37,246	1.51							-1.26
Mar. 25, 1956			33,522	1.36	42	3724	0.15	1956: 1.20	
Aug. 24, 1956	33,127	1.35							-0.16
Apr. 20, 1957			31,100	1.26	33	2027	0.09	1957: 0.96	
July 30, 1957	25,937	1.05							-0.30
Apr. 8, 1958			23,970	0.97	29	1967	0.08	1958: 1.28	
Aug. 24, 1958	33,435	1.36							+0.31
	Total:			16.41				16.50	+0.09
	Annual average:			2.05	42			2.06	+0.01

calculated (Table 1) that the actual ground-water discharge during an average base-flow recession of the Elbow River amounts to 2.05 inches. In other words, 42 per cent of the normal total river flow during the period August to April is generally contributed by ground water.

The difference between the total potential ground-water discharge at the beginning of a ground-water recession and the amount of actual ground-water discharge during that recession will be called 'remaining potential ground-water discharge.' This difference indicates the amount of ground water that is still left in aquifer storage. As will be pointed out later, this volume is of great significance in further calculations.

METHOD OF ESTIMATING TOTAL GROUND-WATER RECHARGE

It can be seen from Figures 5 and 6, as well as from Table 1, that there is a considerable difference between the remaining potential ground-water discharge at the end of any given base-flow recession and the total potential ground-water discharge at the beginning of the next base-flow recession; in other words, water has been added to the ground-water resources. Therefore, this difference is called 'total ground-water recharge' during the period between the end of one base-flow recession and the beginning of the next one. Thus, values of total ground-water recharge are obtained by subtracting the amount of remaining potential ground-water flow at the end of any given ground-water recession from the total potential ground-water discharge at the beginning of the next base-flow recession (Table 1).

Most rainfall in the Calgary area occurs during the period between April and August, with a maximum in June. Indirectly, rainfall rather than total precipitation is considered to be the

TABLE 2. Total Rainfall and Total Recharge, 1951-1958

Year	Rainfall, inches	Total Ground-Water Recharge, inches
1951	19.87	3.81
1952	13.85	3.13
1953	14.99	2.22
1954	16.85	2.62
1955	7.05	1.28
1956	11.52	1.20
1957	9.20	0.96
1958	10.25	1.28
$N = 8$	$\Sigma x = 103.58$	$\Sigma y = 16.50$
Mean	12.94	2.06

chief recharging agent. Data on rainfall over the entire Elbow River basin are not available, but the records for the city of Calgary are believed to be representative of the general weather conditions in the entire basin. Consequently, the correlation between annual rainfall in Calgary and the corresponding total ground-water recharge is relative (Table 2), and the amount of ground-water recharge can only be determined as a basin average.

If figures for total annual rainfall in Calgary are plotted against the corresponding values of total ground-water recharge, a simple linear relationship is suggested (Fig. 7). One of the reasons the points do not all fall on a straight line is that the amount of recharge is determined not only by the *amount* of rainfall but also by the *rain intensity*. The relation between total annual rainfall and total ground-water recharge is given by the regression equation derived from Figure 7:

$$y = 0.21x - 0.76 \quad (4)$$

where y = total ground-water recharge in inches and x = total annual rainfall in inches.

The mean total annual rainfall in Calgary is 11.84 inches, the mean annual rainfall over the entire Elbow River basin is estimated to be 16 inches, and an approximation of the rainfall over the entire basin may be obtained by multiplying the values for the Calgary rainfall by $16/11.84 = 1.35$. The upper scale in Figure 7 shows values thus obtained.

Figure 7 shows that 3.6 inches of rain in Calgary (corresponding to approximately 4.8

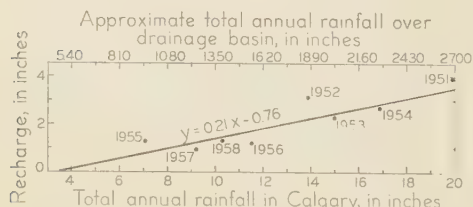


Fig. 7. Relationship between annual rainfall in Calgary and amount of total ground-water recharge.

inches of rain over the entire drainage basin) will produce no recharge at all, probably because sufficient quantities of direct runoff are available to cause the river to be influent.

As has been mentioned in the foregoing paragraph, meteorological records of Calgary show an average total annual rainfall of 11.84 inches, and by use of (4) it can be calculated from this fact that there is an average total ground-water recharge of 1.75 inches. During the period 1950 to 1958 the mean total annual rainfall was 12.94 inches, and the mean total ground-water recharge was 2.06 inches (Table 1). From the nature of the regression equation one sees that these values are in agreement with the relation given in (4).

The total potential ground-water discharge of

any one year is not necessarily the same as the total potential ground-water discharge of the previous or following year (Table 1). This difference between successive values of total potential ground-water discharge is called the 'change in ground-water storage' during the intervening period. The values of changes in ground-water storage, being the third entity in the basic hydrologic equation, are also included in Table 1 and are expressed as depth in inches over the drainage area.

QUANTITATIVE CONSIDERATIONS CONCERNING BASE-FLOW COMPONENTS

Cairn Heights (Fig. 8), which is a small drainage basin west of Calgary, that contributes

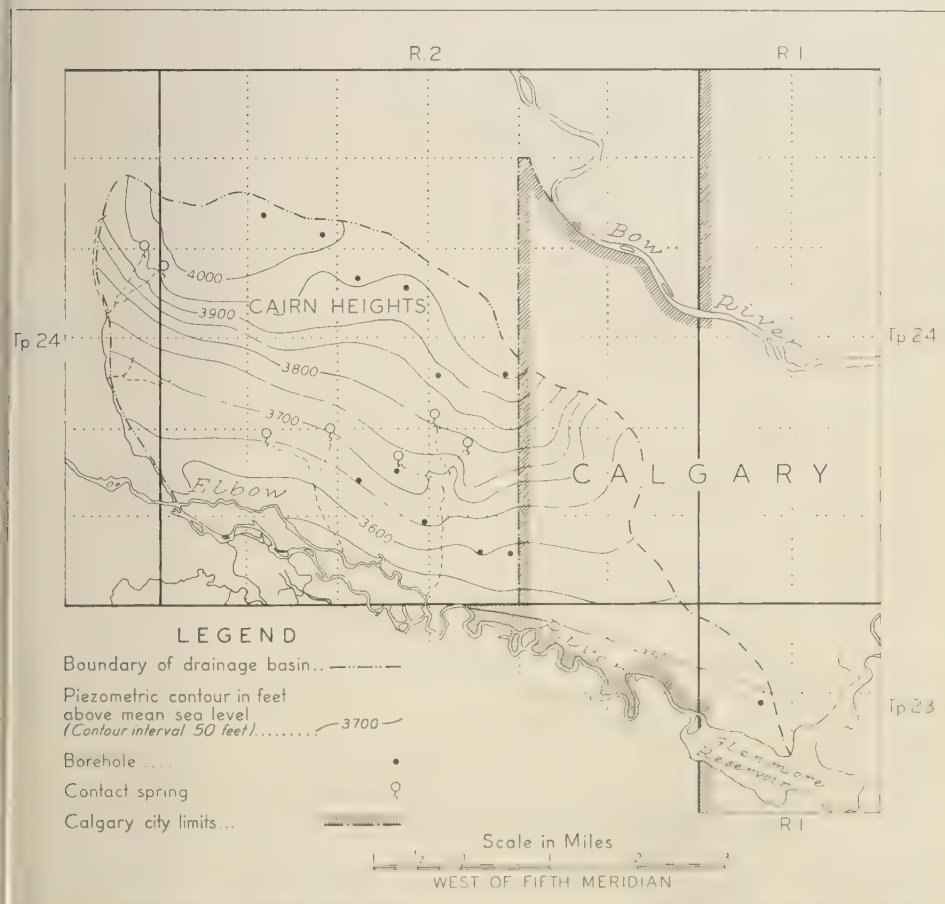


Fig. 8. Map showing piezometric surface of Cairn Heights drainage basin.

to the Elbow River, was selected for further quantitative studies concerning the various base-flow components. The basin comprises 14,138 acres and is in every respect representative of the highlands bordering the Bow and Elbow rivers. The area is covered by an average of 20 feet of fairly permeable drift, underlain by interbedded sandstones and shales of the Paskapoo formation (Fig. 2). A number of contact springs mark the drift-bedrock boundary. The few domestic water wells in the area obtain water from the Paskapoo formation, and the drillers' reports of these wells show transmissibility values ranging between 100 and 200 gpdf. The Elbow River valley along this basin is a half-mile wide, and the stream meanders over and along thick gravel deposits in the valley.

It has been explained in foregoing paragraphs that the average ground-water discharge during a normal base-flow recession amounts to 2.05 inches ($=0.1708$ foot). Accordingly, the Cairn Heights drainage basin should yield 14,138 acres $\times 0.1708$ foot, which is approximately 2400 acre-feet during an average recession lasting 250 days.

According to R. N. Farvolden (1960, personal communication), contact springs along the North Saskatchewan River, near Edmonton, where the drift is of similar thickness, yield approximately 0.37 inches of water per square mile ($=0.0312$

foot). If this value is used in the Calgary area, the yield of contact springs would amount to 14,138 acres $\times 0.0312$ foot, or approximately 450 acre-feet. The yield of contact springs in the Calgary area is probably slightly higher than the yield for those near Edmonton, owing to the coarser drift sediments near Calgary, but from a consideration of their annual yield rather than the yield during the base-flow recession only, this difference is believed to be cancelled. It follows from this calculation that contact springs contribute 19 per cent of the total base-flow recession.

The amount of artesian leakage can be calculated by Darcy's law, which can be expressed as $Q = TIL$, where T is the average transmissibility of the bedrock (150 gallons per day per foot) I is the average gradient toward the river of the piezometric surface (137.5 feet per mile), and L is the length of the strip along which artesian leakage takes place (9.5 miles). Substituting the above values, we obtain $150 \times 137.5 \times 9.5 = 195,937$ gallons per day, which amounts to about 200 acre-feet during an average base-flow recession. This shows that only 8 per cent of the average ground-water recession is contributed by artesian leakage.

The amount of ground water released from bank storage can now be obtained by subtraction; it appears to be about 1750 acre-feet

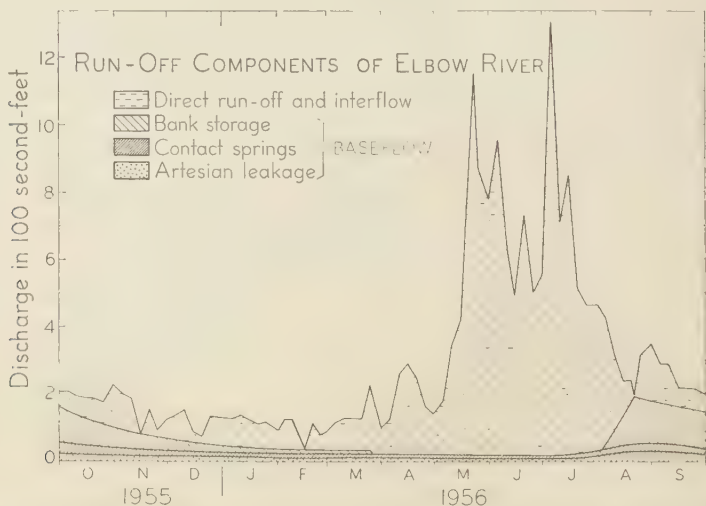


Fig. 9. Hydrograph of Elbow River showing schematically the magnitude of the various base-flow components.

TABLE 3. Amount of Recharge to Base-Flow Components

Base-Flow Components	Amount of Recharge, inches	Percentage of Total Recharge	Percentage of Mean Total Annual Rainfall in Calgary, 1950-1958
Contact springs	0.39	19	3
Artesian leakage	0.16	8	1
Bank storage	1.51	73	12
Total recharge	2.06		16

which is 73 per cent of the actual ground-water discharge during a normal base-flow recession. This amount is released from bank storage along 15 miles of the Elbow River. The river valley along Cairn Heights drainage basin, being 0.5 mile wide, comprises approximately 4900 acres, which means that the average discharge from bank storage represents a depth of 1750 acre-feet/4900 acres = 0.36 foot, spread evenly over the river valley. Knowing that the gravels have a specific yield of approximately 20 per cent, we can expect the ground-water table in the river gravels to show an over-all fluctuation of the order of 2 feet, and the hydrograph of the Calgary Brewery observation well (Fig. 4) indeed shows fluctuations of this order. The magnitude of the various base-flow components is shown schematically in Figure 9.

According to the basic hydrologic equation, ground-water recharge = ground-water discharge \pm change in ground-water storage (5) it is now possible to estimate the amount of recharge to the individual components of base flow. Substituting the average values of actual discharge, total recharge, and change in storage from Table 1 into (5), we obtain the ground-water balance for the Calgary area between 1950 and 1958: 16.50 inches = 16.41 inches \pm 0.09 inches, or mean annual values of 2.06 inches = 1.05 inches \pm 0.01 inches. This balance indicates that, on the average, near-equilibrium exists between recharge and discharge, and it may therefore be assumed that the quantities of recharge to the various base-flow components are proportional to the quantities of discharge. The average recharge to the base-flow com-

ponents has thus been calculated, and the results are presented in Table 3, which also shows the percentage of mean annual rainfall, during 1950-1958, that has been used for ground-water recharge.

CONCLUSIONS

The method of hydrograph analysis outlined is believed to be a potentially useful method of evaluating a regional ground-water balance. It gives quantitative information regarding the entities that make up the basic hydrologic equation. The method is of particular value for areas where stream-gaging records and meteorological data have been collected but where a detailed hydrogeological survey would be uneconomic or otherwise impossible.

The present work suggests that along the eastern edge of the Canadian foothills and Rocky Mountains, where Cretaceous shales and sandstones underlie glacial drift, bedrock and drift aquifers are a very limited source of ground water, owing to the extremely low rate of natural ground-water recharge to these aquifers. Ground water released from bank storage is an important contributor to dry-weather stream flow in this area; consequently, aquifers from which large quantities of water can safely be withdrawn are to be expected only in alluvial deposits of sufficient thickness along the present rivers.

REFERENCES

- Barnes, B. S., The structure of discharge recession curves, *Trans. Am. Geophys. Union*, pt. 4, 721-725, 1939.
- Butler, S. S., *Engineering Hydrology*, Prentice Hall, Inc., Englewood Cliffs, N. J., 356 pp., 1957.
- Foster, J. W., and R. N. Farvolden, A general outline of groundwater conditions in the Alberta Plains Region, *Research Council of Alberta Prelim. Rept. 58-1*, 35 pp., 1958.
- Government of Canada;
 - Dept. of Northern Affairs and Natural Resources; *Surface Water Supply of Canada; Arctic and Western Hudson Bay Drainage*, 1950-1959.
 - Dept. of Transport, 1959, *Annual Meteorological Summary and long-term records, 1885-1959*.
- Hubbert, M. K., The theory of groundwater motion, *J. Geol.*, 58, 785-944, 1940.
- Jones, J. F., Groundwater geology of the Beaverlodge district, Alberta; *Research Council of Alberta Prelim. Rept. 59-2*, 52 pp., 1960.
- Laycock, A. H., Precipitation and stream flow in the Mountain and Foothill region of the Saskatchewan River basin, *Prairie Provinces Water Board, Regina*, 34 pp., 1957.

- Laycock, A. H., Drought patterns in the Canadian prairies, *Assoc. Intern. d'Hydrol. Sci., General Assembly of Helsinki, Publ. 51*, 34-47, 1960.
- Linsley, R. K., M. A. Kohler, and J. L. H. Paulhus, *Applied Hydrology*, McGraw-Hill Book Co., New York, 689 pp., 1949.
- Meinzer, O. E., Outline of groundwater hydrology, *U. S. Geol. Survey, Water Supply Paper 494*, 68 pp., 1923.
- Meyer, O. H., Analysis of runoff characteristics. *Trans. Am. Soc. Civil Engrs.*, 105, 83-89, 1940.
- Natermann, E., Die Linie des langfristigen Grundwassers und die Trockenwetter abflusslinie (TWL), *Die Wasserwirtschaft*, Sonderheft Gewas-
- serkundlichen Tagung in München, p. 12-14, 1950.
- Todd, D. K., Groundwater flow in relation to flooding stream, *Proc. Am. Soc. Civil Engrs.*, 81 sep. 628, 20 pp., 1955.
- Todd, D. K., *Groundwater Hydrology*, John Wiley & Sons, New York, 336 pp., 1959.
- Wenzel, L. K., Several methods of studying fluctuations of ground water levels. *Trans. Am. Geophys. Union*, 400-405, 1936.

(Manuscript received November 14, 1960;
revised January 19, 1961.)

Interrelationships of Watershed Characteristics¹

DON M. GRAY

*Department of Agricultural Engineering
Iowa State University of Science and Technology, Ames, Iowa*

Abstract. The application of the principles of dimensional analysis to obtain the relationships between characteristics of the unit hydrograph and topographic and morphometric properties of a watershed is not possible unless careful consideration is given to the selection of variables. Evidence is presented which shows that, in small watersheds, drainage-area size A , length of the main stream L , and length to center of area L_{ca} are highly correlated. In addition, the results indicate that, when consideration is given to regional influence, the slope of the main stream S_c can be expressed as a function of A , L , or L_{ca} .

Introduction. Unit-hydrograph synthesis for aggregated basins is based on empirical expressions which relate pertinent physical characteristics of a watershed to geometric aspects of the unit hydrograph. These relationships are predicated on the basis that the unit hydrograph of an area represents the integrated effect of all the sensibly constant basin factors and their modifying influence on the translation and storage of a runoff volume from a uniform excess rain occurring during a unit period of time.

Dimensional analysis [Murphy, 1950] has proved to be a useful tool in engineering fields for establishing relationships within a system of variables. Strahler [1958] discussed the application of these principles to the field of geomorphology in connection with fluvially eroded landforms. These studies give rise to the thought that the principles of dimensional analysis may be used to derive the desired relationships between watershed characteristics and unit-graph properties required for hydrograph synthesis.

A study of this type was made by Gray [1960], who used hydrologic and topographic data collected from 42 small watersheds located throughout the United States. The results of the study indicated, however, that the approach used was actively unsuccessful. Consequently, an additional investigation was undertaken to determine plausible reason for this failure.

According to the theory of dimensional analysis, the variables must be selected with great care so that the π terms formed in their combination

will remain independent. It was ascertained that failure in the initial study may have resulted from violation of this criterion by the selection of nonindependent watershed variables. The results of a subsequent study, reported herein, substantiate this presumption. Evidence that many watershed characteristics are highly correlated is presented.

Theory. The field of quantitative geomorphology has received considerable attention in recent years. These studies were stimulated by the work of Horton [1945], who suggested that the development of morphological characteristics depends on three main factors: surface resistivity to sheet erosion, runoff intensity, and ground slope. Since these factors vary with soil and bedrock conditions, vegetative cover, and climatic conditions, watersheds in different regions would be expected to exhibit wide differences in the degree of development of their drainage systems. Miller [1953, p. 12] explained the differences in the morphological character of watersheds attributable to lithological differences by comparing an area of sandy soil underlain by porous and permeable sandstone with an area of dense clayey soils underlain by dense shale: 'Maturely dissected topography would be expected to vary in the following ways: in sandstone areas streams of a given order should be longer and their drainage basins larger; drainage density should be less, slopes should be longer and possibly gentler than in the shale areas.' Furthermore, in a dolomite area the characteristics would be expected to differ from those in either of the areas described above.

Horton [1945] proposed three laws related to

¹Journal Paper No. J-4014 of the Iowa Agricultural and Home Economics Experiment Station, Ames, Iowa. Project No. 1266.

fluviially eroded streambeds. These are given in mathematical form by the following expressions.

Law of stream numbers:

$$N_o = r_b^{s-o} \quad (1a)$$

where

N_o = number of streams of order o .

r_b = bifurcation ratio, the ratio of the number of stream segments of given order N_o to the number of stream segments of the next higher order N_{o+1} ; i.e., $r_b = N_o/N_{o+1}$.

s = stream segments of highest order in the watershed.

Law of stream lengths:

$$l_o = l_1 r_l^{o-1} \quad (1b)$$

where

l_o = average length of streams of order o .

l_1 = average length of first-order streams.

r_l = stream length ratio or the ratio of the average length of streams of given order

l_o to the average length of streams of the next lower order l_{o-1} ; i.e., $r_l = l_o/l_{o-1}$.

Law of stream slopes:

$$s_o = s_1/r_s^{o-1} \quad (1c)$$

where

s_o = average slope of the streams of order o .

s_1 = average slope of first-order streams.

r_s = slope ratio or the ratio of the average slope of streams of given order s_o to the average slope of streams of the next higher order s_{o+1} ; i.e., $r_s = s_o/s_{o+1}$.

Schumm [1956, p. 606] proposed a fourth law of drainage composition, the law of stream area as a sequel to Horton's laws:

$$a_o = a_1 r_a^{o-1} \quad (1d)$$

where

a_o = mean drainage area of streams of order o .

a_1 = mean drainage area of first-order basin.

r_a = area ratio, ratio of average basin area of stream of given order a_o to the average area of basins of the next lower order a_{o-1} ; i.e., $r_a = a_o/a_{o-1}$.

According to these laws, N_o , l_o , s_o , and a_o are related either by a direct or an inverse geometric series to stream order. It follows that the interrelationships between these variables for watersheds of given order can be expressed by a series of simple power equations.

Topographic data. Topographic data were collected from 47 small watersheds in Illinois, Iowa, Missouri, Nebraska, North Carolina, Ohio, and Wisconsin. The watersheds varied in size from 0.23 to 33.00 square miles. A list of the pertinent topographic and morphometric properties is given in Table 1 and Figure 1.

The type of topographic information used to derive these properties varied considerably. For the watersheds in Missouri, U. S. Geological Survey (USGS) quadrangle topographic maps with scales of 1:62,500 or 1:24,000 and contour intervals of 20 feet were generally available. Properties of the Ohio watersheds were taken from USGS topographic maps with a scale of 1:4800 and contour intervals of 5 feet and those for the North Carolina watersheds from U. S. Forest Service maps with a scale of 1:7920 and contour intervals of 50 feet. For the watersheds

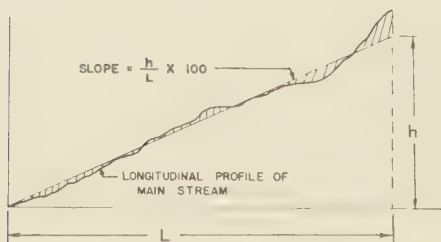
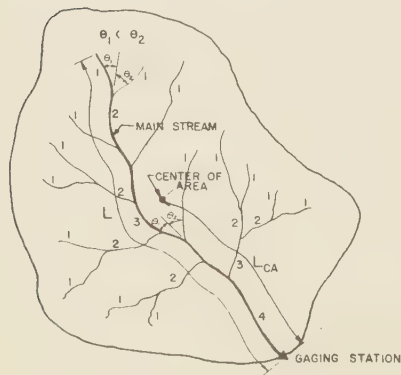


Fig. 1. Watershed characteristics.

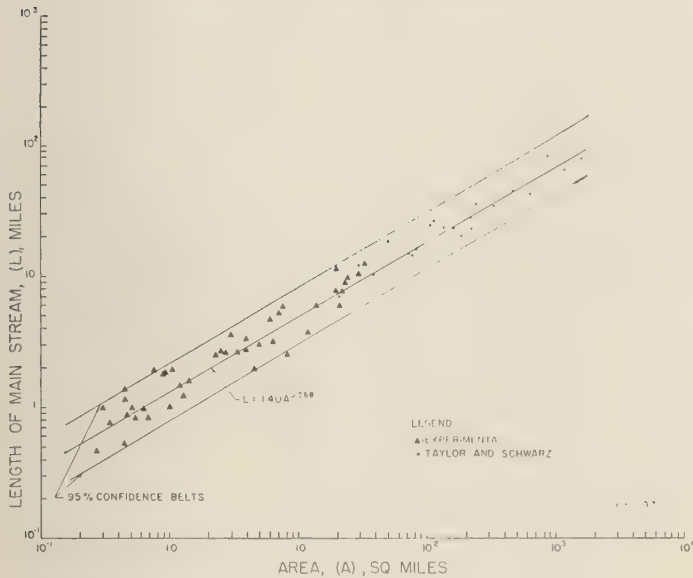


Fig. 2. Relation of length of main stream L and watershed area A .

Illinois, Nebraska, and Wisconsin, operated by the Agricultural Research Service, topographic maps with a scale of 1:2400 and contour intervals of 5 feet were available.

Considerable difficulty was encountered in obtaining information for watersheds in Iowa and Nebraska. In these states, contour maps were obtainable for only three of the 13 watersheds. The length and shape characteristics of the other 10 basins were obtained from drainage maps having various scales whereas the stream gradients were taken with a barometric altimeter during on-site visits to each of the watersheds. Length of main stream L and drainage-basin area A . By combining (1b) and (2) it can be shown that within a homogeneous watershed the average length of streams l_o from basins of given order are related to the mean size of basin a_o by the expression

$$l_o = l_1(a_o/a_1)^z \quad (3)$$

where

- l_1 = average length of first-order streams.
- a_1 = mean area of first-order basins.
- z = ratio, $\ln r_1 / \ln r_a$, where r_1 and r_a are the length and area ratios, respectively.

Since the slope z is positive, (3) implies that l_o increases as a power series with a_o . Morisawa

[1959] has shown this property to exist in watersheds in the Allegheny Plateau.

In this study, no attempt was made to categorize stream order because of the errors that would be introduced by using the type of topographic information available [Coates, 1958]. The length of main stream L , as defined herein, represents a composite of stream lengths of different order. However, since, under comparable conditions, larger watersheds sustain longer streams than smaller watersheds, it would be expected that L would vary with A .

The parameters L and A for the 47 watersheds were plotted on logarithmic paper (Fig. 2). These data were supplemented with comparable results for watersheds in the North and Middle Atlantic States [Taylor and Schwarz, 1952]. The regression line fitted to the data by the method of least squares is defined by the equation

$$L = 1.10 A^{0.568} \quad (4)$$

where L is in miles and A in square miles. The standard error of estimate was computed to be 24.8 per cent.

The practical importance of the regression is that it provides the evidence necessary to show that the variables L and A are not independent. Thus, their use as independent terms in dimen-

TABLE 1. Topographic Data

State	Number	Watershed	A , [*] mi ²	L , [†] mi	L_{ca} , [‡] mi	R_e , [§]	R_c ,	S_c , [¶] %
Illinois	1	W-IV, Edwardsville	0.45	0.54	0.28	0.88	0.76	1.10
Iowa	2	Davids Creek near Hamlin	26.01	9.14	4.95	0.70	0.66	0.39
	3	Hayworth main outlet near Climbing Hill	0.91	1.80	0.85	0.55	0.53	1.41
	4	Indian Creek at Council Bluffs	7.56	5.69	2.08	0.53	0.45	0.49
	5	Muckey Creek near Mapleton	0.69	0.83	0.45	0.68	0.50	1.34
	6	Nepper main outlet near Mapleton	0.35	0.75	0.43	0.77	0.80	2.56
	7	Ralston Creek near Iowa City	3.00	3.50	2.80	0.48	0.47	0.45
	8	Rapid Creek near Iowa City	24.57	9.50	4.15	0.56	0.56	0.21
	9	Renneker Main Outlet near Anthon	0.89	1.78	0.68	0.66	0.60	0.94
	10	Waubonsie Creek near Bartlett	32.64	12.50	5.30	0.53	0.39	0.40
Missouri	11	Beaver Creek near Rolla	13.70	5.95	2.90	0.67	0.46	0.70
	12	Behmke Branch near Rolla	1.03	1.95	1.15	0.57	0.60	1.37
	13	Big Creek near Yukon	8.36	2.45	1.65	0.96	0.81	2.28
	14	Bourbeuse Creek near St. James	21.30	6.00	3.02	0.84	0.69	0.41
	15	Coyle Branch at Houston	1.30	1.21	0.80	0.75	0.86	0.52
	16	East Fork Fishing River at Excelsior Springs	20.00	7.80	3.60	0.63	0.70	0.50
	17	Green Acre Branch near Rolla	0.62	0.98	0.60	0.90	0.87	1.45
	18	Jenkins Branch at Gower	2.72	2.50	1.20	0.68	0.70	0.53
	19	Lanes Fork near Rolla	0.23					
	20	Lanes Fork near Vichy	24.10	9.40	3.10	0.58	0.39	0.40
	21	Little Beaver Creek near Rolla	6.27	3.10	1.60	0.63	0.70	1.02
	22	Lost Creek at Elsberry	12.20	3.70	1.98	0.88	0.65	0.74
	23	Mill Creek at Oregon	4.90	3.00	1.60	0.81	0.68	0.79
	24	Oak Grove Branch near Brighton	1.00	1.00	0.75			
	25	Shiloh Branch near Marshall	2.87	2.45	1.60	0.54	0.64	0.45
	26	Stahl Creek near Miller	3.86	2.70	1.40	0.74	0.60	0.59
	27	Stark's Creek at Preston	4.72	1.98	1.10	0.62	0.84	
	28	White Cloud Creek near Maryville	6.06	4.60	2.60	0.46	0.42	
Nebraska	29	Dry Creek near Curtis	20.00	11.59	5.49	0.43	0.33	0.30
	30	W-3, Hastings	0.75	1.96	1.54	0.68	0.57	0.86
	31	New York Creek near Herman	30.00	10.25	5.45	0.53	0.68	0.25
	32	Tekamah Creek at Tekamah	21.53	7.50	4.25	0.65	0.38	0.52

* Plane area of the watershed which is enclosed within the topographic divide above the gaging station.

† Distance along the main stream from the gaging station to the outermost point defined on the topographic map (Fig. 1). The main stream is the stream of highest order which passes through the gaging station. To delineate the main stream at bifurcations, the following rules suggested by Horton [1945, p. 28] were used:

1. Starting below the junction, the main stream was projected upstream from the bifurcation in the same direction. The stream joining the main stream at the greatest angle was taken as the lower order (Fig. 1).

2. If both streams were at about the same angle to the main stream at the junction, the shorter was taken as the lower order.

‡ Distance along the main stream from the gaging station to the point nearest the mass center of area (Fig. 1).

§ Elongate ratio.

|| Circularity ratio.

¶ Slope of a line drawn along the longitudinal section of the main channel in such a manner that the area between the line and a horizontal line drawn through the channel outlet elevation is equal to the area between the channel grade line and the same horizontal line (Fig. 1).

TABLE 1. Continued

State	Number	Watershed	A ,* mi ²	L ,† mi	L_{ca} ,‡ mi	R_e ,§	R_c ,	S_c ,¶ %
Ohio	33	W-5, Coshocton	0.55	0.82	0.48	0.84	0.77	2.64
	34	W-11, Coshocton	0.46	1.17	0.74	0.60	0.53	1.83
	35	W-91, Coshocton	0.46	1.31	0.57	0.60	0.59	2.13
	36	W-92, Coshocton	1.44	1.56	0.72	0.80	0.70	1.84
	37	W-94, Coshocton	2.37	2.41	1.02	0.73	0.68	1.37
	38	W-95, Coshocton	4.02	3.25	1.45	0.68	0.53	1.09
	39	W-97, Coshocton	7.15	5.11	2.44	0.58	0.54	0.72
	40	W-196, Coshocton	0.47	0.88	0.42	0.60	0.35	3.94
Pennsylvania	41	W-I, Fennimore	0.52	0.99	0.55	0.87	0.58	1.80
	42	W-IV, Fennimore	0.27	0.46	0.29	0.98	0.75	2.20
North Carolina	43	7, Coweeta	0.23	0.59	0.41	0.78	0.68	19.90
	44	8, Coweeta	2.93	2.38	1.34	0.68	0.65	9.02
	45	9, Coweeta	2.79	2.76	1.53	0.62	0.57	8.50
	46	10, Coweeta	0.33	0.95	0.60	0.65	0.56	19.10
	47	4, Bent Creek	1.21	1.45	0.94	0.66	0.51	
						Mean	0.68	0.61
						Coefficient of variation	0.20	0.23

nal-analysis techniques is prohibited. In addition, (4) can be used to obtain an approximation of the length of the main stream once the drainage-basin size is known.

The 95 per cent confidence belts have been added to Figure 2 to facilitate the use of the equation as a prediction equation. Similarly, a measure of the degree of association between the variables is given by the correlation coefficient r . Length to center of area L_{ca} and length of main stream L . The agreement of L and A with common regression suggests that the shapes of the watersheds may not differ appreciably. In hydrologic work, the length to center of area is frequently taken as a quantitative measure of basin shape. This parameter is shown plotted with L in Figure 3. As before, selected data reported by Taylor and Schwarz are included.

The relation between L_{ca} and L is defined by the equation

$$L_{ca} = 0.54L^{0.98} \quad (5a)$$

For this regression, the standard error of estimate is determined to be 14.8 per cent.

The result implies (1) that the interdependence of the parameters L_{ca} and L prohibits their use in independent terms and (2) that the use of the product term LL_{ca} , as is frequently done in hydrologic studies, has little advantage over the use of L or L_{ca} alone. In practical work, only

small error will be introduced if the relationship is applied as

$$L_{ca} = 0.50L \quad (5b)$$

Discussion. The compactness or shape characteristics of the watersheds were compared with those of 340 drainage basins from the north-eastern United States reported by Langbein and others [1947]. They evaluated the area-distance property for each of the watersheds by the

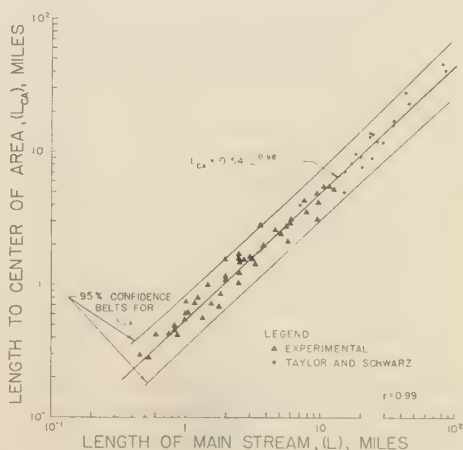


Fig. 3. Relation of length to center of area L_{ca} and length of main stream L .

factor $\sum al$ or the product of each partial area a by l , the channel distance from the midpoint of the main stream to the*downstream gaging station. The regression of $\sum al$ with drainage area A for the 340 drainage basins was determined to be

$$\sum al = 0.90 A^{1.56} \quad (6a)$$

By definition, $L_{ca} = \sum al/A$; thus (6a) may be written

$$L_{ca} = 0.90 A^{0.56} \quad (6b)$$

The properties of the watersheds investigated in the current study can be expressed in comparable form by combining (4) and (5a) to obtain

$$L_{ca} = 0.74 A^{0.55} \quad (7)$$

Equations 6b and 7 define two lines having practically the same slope but different intercept values. This property is taken to be an indication that the watersheds studied were generally more compact than those reported by Langbein and others.

For illustrative purposes, the equational forms of a few simple geometrical shapes were considered (Table 2).

The values of the exponent and coefficient of (7) differ from those of any of the geometric forms considered. A review of the topographic data showed that the general shape of the watersheds appeared to fall between ovoid and pear shape. This observation does not, however, contradict the result given by (5b), which also defines a square, a rectangle, or a circle. Because of the sinuosity of the streams, L for the watersheds is longer than a comparable dimension for any of the above forms.

Two other shape factors were also computed;

TABLE 2. Relationship between Basin Characteristics L_{ca} and A for Three Simple Geometric Forms [Langbein and others, 1947, p. 135]

Geometric Shape	Equational Form between L_{ca} and A
Circle with outlet at center, 'Glory hole'	$L_{ca} = 0.375A^{0.50}$
Equilateral triangle with outlet at one of the vertices	$L_{ca} = 0.94A^{0.50}$
Square with outlet at one of the corners	$L_{ca} = 0.76A^{0.50}$

the circularity ratio [Miller, 1953, p. 8] and the elongate ratio [Schumm, 1956, p. 612]. These ratios are defined as follows.

Circularity ratio:

$$R_c = A/A_c$$

where

R_c = circularity ratio.

A = basin area.

A_c = area of a circle with equal perimeter to the basin.

Elongate ratio:

$$R_e = D/L'$$

where

R_e = elongate ratio.

D = diameter of a circle with equal area to the basin.

L' = maximum length of the basin parallel to the principal drainage lines.

The maximum value the ratios can attain unity, for a circular watershed.

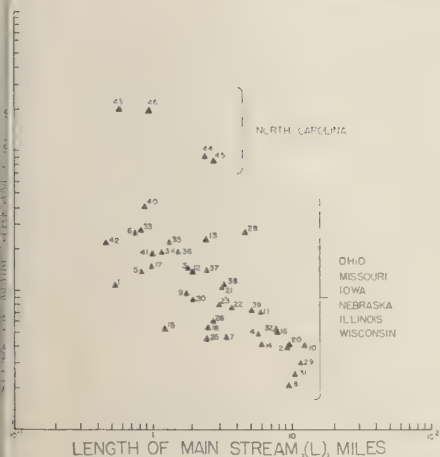
Miller [1953] and Coates [1958] have reported that, unlike stream length and basin area, the circularity of basins remains reasonably constant regardless of differences in lithological composition. Miller suggested further that the constant property of the R_c values may be an expression of a universal equilibrium form. For the watersheds investigated, the mean R_c value was computed to be 0.61 and to have a coefficient of variation equal to 23 per cent. The magnitude of the coefficient of variation compares favorably with values listed by Melton [1958] and is indicative of the relative stability of the parameter R_c .

Similarly, the mean and the coefficient of variation of the elongate ratios were computed to be 0.68 and 20 per cent, respectively. The results suggest that this parameter is perhaps a more stable property of basin geometry than the circularity ratio.

Length of main stream L and channel gradient S . Horton's laws of stream lengths and stream slopes (see equations 1b and 1c) can be combined in the form

$$s_o = s_1(l_1/l_o)^b$$

where b , the slope on logarithmic paper, is the ratio $\ln r_s / \ln r_e$. The relation is valid provided



4. Variation of the slope of the main stream S_e with length L .

area is homogeneous and s_e and l_e follow the patterns established previously. Frequently, however, the proportion given by (8) is inaccurate because of changing soil and lithological conditions along the stream or changes brought about by artificial or natural controls.

For the watersheds studied, the average slope of the main stream S_e was plotted with the length L (Fig. 4). As is shown in the figure, the watersheds exhibit marked differences in their type characteristics; the most outstanding difference is the extremely steep gradients of the North Carolina watersheds.

Such wide variation is not unexpected, however. Unlike the form elements which involve only a two-dimensional measure of erosional development, when stream gradients are considered the problem becomes three-dimensional, and surface resistivity exerts a predominant effect. Among the watersheds studied, in addition to marked climatic, vegetative, and ground-slope differences, there are contrasting soil and bedrock conditions. For example, the residual soils of the Ohio watersheds, which have a shallow solum, are moderately permeable and cover a shale or slate parent material [U. S. Dept. Agr., 1941]. In contrast, watersheds within the Nebraska-western Iowa region have thick, silty, highly permeable, loessial soils [Riecken and Smith, 1949]. Thus, on similar watersheds within the two regions, the runoff volume would be greater in Ohio and erosion would be more severe in

Nebraska and western Iowa. Likewise, watersheds in Nebraska and western Iowa have flatter gradients and lower drainage densities than those in Ohio.

The variables S_e and L from watersheds within three regional groups were plotted separately as shown in Figure 5. The data can be interpreted by using the results obtained from an analysis of covariance. This test gives statistical confirmation (1) that the slopes of the regression lines are not significantly different and (2) that the differences between the adjusted mean values of the three groups is greater than that which can be accounted for by sampling variation. In effect, the data could be represented by three parallel lines drawn through the mean logarithmic values of S_e and L of each group. The high r values listed (Fig. 5) are also indicative of the high degree of association between the variables when consideration is given to regional influence.

Figure 4 shows that the plotted data for the other watersheds in Illinois, central Iowa, Missouri, and Wisconsin have no general pattern but vary appreciably in their relative positions. In some cases they approach the regression line for the Ohio watersheds; in others they approach that of the Nebraska-western Iowa region or appear to occur in their own individual class. Since the characteristics of these basins were not available, the delineation of a complete family of curves was not attempted.

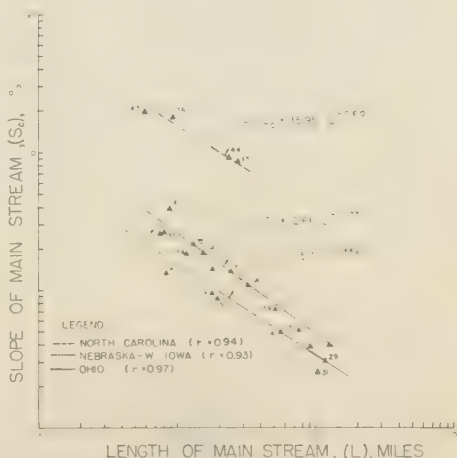


Fig. 5. Relation of slope of the main stream S_e and length of the main stream L for watersheds within selected regional groupings.

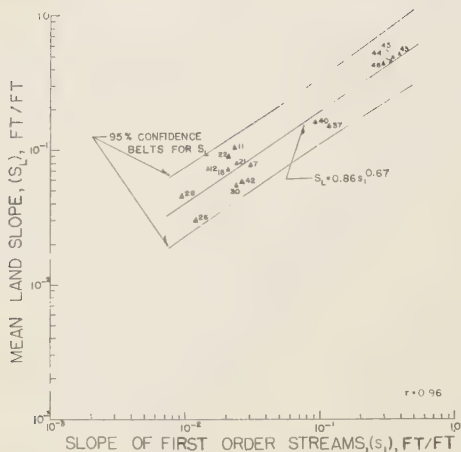


Fig. 6. Relation of mean land slope S_L and average slope of a representative number of first-order streams from the same watershed s_1 .

Mean land slope S_L and average slope of first-order streams s_1 . In hydrologic investigations involving small watersheds, some quantitative measure of the general land slope is usually desired. One frequently used measure is the mean land slope S_L , which is evaluated by either the intersection-line method or the grid-intersection method [Horton, 1945]. However, the labor involved in computation with either of these methods is extensive, and topographic maps are required.

To overcome these limitations, an attempt was made to relate S_L to a more readily measurable characteristic of a watershed. Figure 6 shows the mean land slopes S_L of 16 watersheds, determined by the grid-intersection method and plotted with the mean slope obtained by averaging the gradients of a representative sample of first-order streams s_1 taken from the same basins. The regression of these data is defined by the equation

$$S_L = 0.86s_1^{0.67} \quad (9)$$

A standard error of estimate for the S_L values was calculated to be 28.6 per cent.

Equation 9 furnishes a simple relation from which an estimate of the mean land slope can be obtained from the slopes of the first-order streams. The magnitude of s_1 can easily be derived from readings taken with a barometric altimeter or from information taken directly

from topographic maps. When topographic maps are used, first-order streams should be defined by the contour method [Morisawa, 1957].

Summary and conclusions. The results of this study indicate that the application of the principles of dimensional analysis to assist in developing relationships useful for hydrograph synthesis is not feasible unless careful consideration is given to the selection of watershed parameters. In an analysis of several topographic characteristics from a number of small watersheds of different vegetative, soil, lithologic, physiographic, and climatic conditions, the following conclusions were reached:

1. The length of the main stream L , drainage basin area A , and length to the center of area L_{ca} are highly correlated.
2. The general geometric shape of small watersheds falls between ovoid and pear shape.
3. The slope of the main stream S_c can be inversely related to the parameters L , L_{ca} , and A as a simple power equation if consideration is given to regional influence.

REFERENCES

- Coates, D. R., Quantitative geomorphology of small drainage basins of Southern Indiana, *Tech. Rept. 10*, Dept. of Geol., Columbia Univ., New York, 1958.
- Gray, D. M., Derivation of hydrographs for small watersheds from measurable physical characteristics, Unpublished Ph.D. Thesis, Library, Iowa State Univ. Sci. and Technol., Ames, Iowa, 1957.
- Horton, R. E., Erosional development of streams and their drainage basins; hydrophysical approach to quantitative morphology, *Bull. Geol. Soc. Am.*, 56, 275-370, 1945.
- Langbein, W. B., and others, Topographic characteristics of drainage basins, *U. S. Geol. Surv. Water Supply Paper 968-C*, 125-155, 1947.
- Melton, M. A., List of sample parameters of quantitative properties of landforms: their use in determining the size of geomorphic experiments, *Tech. Rept. 16*, Dept. of Geol., Columbia Univ., New York, 1958.
- Miller, V. C., A quantitative geomorphic study of drainage basin characteristics in the Clinch Mountain area Virginia and Tennessee, *Tech. Rept. 3*, Dept. of Geol., Columbia Univ., New York, 1953.
- Morisawa, M. E., Accuracy of determination of stream lengths from topographic maps, *Trans. Am. Geophys. Union*, 38, 86-88, 1957.
- Morisawa, M. E., Relation of quantitative geomorphology to stream flow in representative watersheds of the Appalachian Plateau province, *Tech. Rept. 20*, Dept. of Geol., Columbia Univ., New York, 1959.

- raphy, G., *Similitude in Engineering*, Roland Press, New York, 1950.
- cken, F. F., and G. D. Smith, Principal upland hills of Iowa, *Iowa State College Extension Serv., Agron.* 49 (rev.), mimeo., 1949.
- umm, S. A., Evolution of drainage systems and slopes in badlands at Perth Amboy, N. J., *Bull. Geol. Soc. Am.*, 67, 597-646, 1956.
- decor, G. W., *Statistical Methods; Applied to Experiments in Agriculture and Biology*, 5th ed., Iowa State College Press, Ames, Iowa, 1957.
- ahler, A. N., Dimensional analysis applied to fluvially eroded landforms, *Bull. Geol. Soc. Am.*, 69, 279-300, 1958.
- Taylor, A. B., and H. E. Schwarz, Unit hydrograph lag and peak flow related to basin characteristics, *Trans. Am. Geophys. Union*, 33, 235-246, 1952.
- U. S. Dept. of Agriculture, Hydrologic data, North Appalachian Experimental Watershed, Coshoc-ton, Ohio, *U. S. Dept. Agr., Soil Conserv. Serv. Hydrologic Bull.* 1, 1941.

(Manuscript received December 17, 1960.)

Erodibility of Some California Wildland Soils Related to Their Metallic Cation Exchange Capacity

JAMES R. WALLIS

*Pacific Southwest Forest and Range Experiment Station
Forest Service, U. S. Department of Agriculture, Berkeley, California*

AND

LEE J. STEVAN

University of California, Berkeley, California

Abstract. The inherent erodibility of twenty soils was indexed using Middleton's dispersion ratio and Anderson's surface-aggregation ratio. These indices were then used as dependent variables for several regression analyses. The milliequivalents per 100 grams of oven-dried soil for the four most plentiful soil cations (Ca, Mg, K, and Na) were determined and then used as independent variables in the regressions.

The best fit to the data was found to be an equation of the type:

$$\text{Erosion index} = a + b \left(\frac{\text{meq of}}{\text{Ca}^{++} + \text{Mg}^{++}} \right) + c \left(\frac{\text{meq of}}{\text{Ca}^{++} + \text{Mg}^{++}} \right)^2$$

Both the linear and the curvilinear terms were significant at the arbitrarily selected 5 per cent level.

Introduction. Good indices of inherent soil erodibility have been developed by Middleton [1930] and Anderson [1954]. Middleton's dispersion ratio (D.R.) is a measure of the increase in small particles in a soil sample after work has been performed on it. The dispersion ratio has been found to be a statistically significant indicator of inherent soil erodibility in southern California [Anderson, 1951] and in Oregon [Anderson, 1954]. An excellent experimental study by Smerdon and Beasley [1959] has proved that the dispersion ratio is a good measure of the inherent erodibility of a soil. Anderson's surface aggregation ratio (S/A) has also been related to sediment discharge from watersheds [Anderson, 1954] and has the additional advantage of being highly correlated with the geologic parent material. In this paper a study is reported which showed that these soil erodibility indices are related to amount and kind of cations present in the soil.

One of the cohesive properties of a soil that should aid it in resisting erosive forces is the adsorbed metallic cation complex. Skula and Nayar [1943] have shown that soil permeability can be altered by changing the amount and kind of cations present in a soil. It seems reason-

able to expect that similar relationships could be shown to exist between the amount and kind of cations and the degree of particle aggregation in a soil. If so, the soil erodibility might be changed by modifying the cation concentrations.

Two approaches to this hypothesis are possible: first, the replacement of exchangeable cations and the measurement of the resulting changes in erodibility; second, a regression analysis of the erodibility and the amounts of cations present in different soils. In this study the second of these two approaches was used.

A study of this kind has been previously reported by Rost and Rowles [1941]. They found positive correlations between aggregation and organic matter, and between aggregation and total cation exchange capacity (defined as exchangeable H^+ , Mg^{++} , and Ca^{++}). Between aggregation and Mg^{++} they found no relation, and between aggregation and Ca^{++} or Ca^{++} plus Mg^{++} they found a slight negative correlation. They suggested that the implied positive correlation between aggregation and H^+ was probably caused by the indirect effect of H^+ in retaining organic matter. The soils they tested contained much more organic matter than is found in California's wildland soils. A com-

TABLE 1. Summary Sheet for 20 California Wildland Soils and Their Physical Characteristics

Sample Number	Parent Material	Zone	Veg.	Elev. mm	Percentage of Entire Soil in Class					Clay	Fine Clay	Surface Aggregation Ratio	Dispersion Ratio
					25 mm	5 to 5 mm	5 to 2 mm	Sand 2 to .05 mm	Silt				
9	Granite	N. Coast	Pine	2400	2	4.2	17.1	50.3	17.2	3.8	5.4	144.0	63.6
16	Granite	Sierra	Pine	3600	0	1.4	9.6	55.7	19.9	5.3	8.0	187.8	76.5
41	Basalt	N. Coast	Pine	2050	5	4.1	13.7	29.3	29.3	4.6	13.9	39.1	55.9
43	Basalt	N. Coast	Grass	3000	15	11.6	13.3	25.8	23.4	4.2	6.6	43.6	49.1
42	Basalt	N. Coast	Grass	4300	15	11.5	12.6	31.3	19.8	3.0	6.7	59.9	49.5
63	Serp.	N. Coast	Pine	3400	5	14.8	12.1	29.3	25.2	4.1	9.5	33.6	35.6
64	Serp.	N. Coast	Pine	3700	40	12.7	6.9	17.4	14.9	2.8	5.1	44.4	48.7
66	Serp.	N. Coast	Pine	2700	10	29.7	21.8	22.6	9.8	1.9	4.2	98.7	49.1
67	Serp.	N. Coast	Pine	3900	50	17.2	9.2	11.7	6.9	1.2	3.8	49.7	31.9
103	Schist	Sierra	Grass	2400	35	3.9	1.8	35.2	20.5	0.6	3.0	248.4	84.6
117	Schist	N. Coast	Forest	4000	20	20.8	13.8	22.4	14.8	3.6	4.5	67.9	55.4
118	Schist	N. Coast	Forest	4200	10	29.7	23.2	21.9	9.7	2.6	3.0	79.8	38.6
120	H. Sed.	N. Coast	Pine	3100	0	25.2	28.5	31.6	7.7	1.9	5.1	116.9	40.8
123	H. Sed.	N. Coast	Pine	2800	15	16.6	14.5	19.7	27.8	2.7	3.8	40.3	55.1
124	H. Sed.	N. Coast	Pine	3200	0	6.2	20.5	28.6	25.6	5.1	13.9	33.1	41.7
126	H. Sed.	N. Coast	Pine	4100	0	8.8	19.1	36.0	26.6	3.6	5.8	57.9	50.0
169	Alluvium	N. Coast	Grass	2300	15	21.1	5.4	16.4	24.6	5.3	12.3	14.6	21.1
170	Alluvium	N. Coast	Grass	2800	40	14.0	7.6	26.2	9.1	1.2	1.9	98.4	37.7
171	Alluvium	N. Coast	Grass	2800	0	20.6	23.2	33.9	13.8	3.9	4.5	102.4	54.5
172	Alluvium	N. Coast	Pine	2300	15	18.6	11.5	25.8	18.1	5.5	5.5	39.5	34.0
Range of values				[2050 [4300	0 50	1.4 29.7	1.8 28.5	11.7 55.7	6.9 29.3	0.6 5.5	1.9 13.9	14.6 248.4	21.1 84.6]
Mean of values				3150	15	14.6	14.3	28.6	18.2	3.3	6.2	79.75	48.67

parison between the results of their study and ours would be easier if they had tested the bare mineral soil.

Soils tested in this study. The soil samples used in our study were collected by J. André from mineral soil at depths of 0 to 6 inches, under conditions of constant slope, rainfall, and full vegetative cover for each of eight geologic types. André performed a complete mechanical analysis of 168 soil samples and determined the dispersion and surface aggregation ratios.

We picked 20 of André's samples to give a wide variation in geographic location and geologic type. Relevant information for the 20 soils used is contained in Table 1. (André's samples were taken throughout the northern California Sierra and Coast ranges in areas for which no soil maps are available. No attempt has been made to classify these soils by series. The exact location of each sampling point and other descriptive material are on file at the Pacific Southwest Forest and Range Experiment Station.)

Summary of methods used. Ten-gram samples of each soil were digested on a steam bath for 30 minutes in 40 ml of 1 *N* NH_4OAc (pH7), then leached through a Buechner funnel with Whatman No. 30 filter paper, a procedure similar to that described by Jackson [1958].

Combined exchangeable magnesium and calcium was determined by titration and back titration with standard 0.1 *N* EDTA (disodium hydrogen ethylene-diaminetetraacetic acid, I.W. 372) and standard 0.1 *N* calcium solution with eriochrome black T for an indicator. A separate titration with a calcein indicator determined the amount of calcium and hence, by difference, the amount of magnesium. These titrations were buffered with triethanolamine to prevent iron or manganese from interfering with the indicator, and with sodium cyanide to complex any copper or zinc. Excess EDTA was added to dissolve calcium and magnesium phosphates, and this excess of EDTA made the back titrations with standard calcium necessary. The titrations were corrected for contamination by the use of blanks.

Potassium and sodium were determined with Beckman flame photometer [Jackson, 1958]. The amounts of each of the four metallic cations, Mg, K, and Na present in the samples were expressed on the basis of milliequivalents per

TABLE 2. Amount of Exchangeable Cations in 20 California Wildland Soils and Their Chemical Characteristics

Sample Number	pH	Amount of Exchangeable Cations, milliequivalents per 100 grams of oven-dried soil			
		Ca ⁺⁺	Mg ⁺⁺	K ⁺	Na ⁺
9	6.4	6.89	1.65	0.943	0.052
16	5.7	4.69	1.72	0.382	0.040
41	6.5	7.06	3.68	0.852	0.062
43	6.8	11.59	2.33	0.838	0.063
42	7.3	14.02	1.99	1.242	0.074
63	6.5	14.73	11.02	0.312	0.064
64	6.5	2.09	12.03	0.167	0.072
66	6.9	2.01	22.57	0.189	0.054
67	6.5	7.82	20.46	0.453	0.074
103	6.3	4.73	1.45	0.421	0.083
117	5.9	7.66	2.42	0.476	0.059
118	6.4	18.10	3.36	0.815	0.086
120	6.0	5.85	2.19	0.437	0.067
123	5.7	8.85	4.30	0.350	0.074
124	6.6	7.64	9.32	0.740	0.056
126	6.2	6.54	2.15	0.580	0.067
169	6.1	18.30	7.65	0.583	0.113
170	5.8	15.13	3.32	0.375	0.114
171	5.9	9.16	2.72	0.621	0.056
172	6.1	8.56	25.09	0.674	0.070
Range		{ 2.01 18.30	1.45 25.09	0.167 1.242	0.040 0.114
Mean		9.07	7.07	0.573	0.070

100 grams of oven-dried soil are given in Table 2.

(A fuller explanation of the chemical procedure used in the present study is available upon request to the Division of Watershed Management Research, Pacific Southwest Forest and Range Experiment Station, P. O. Box 245, Berkeley 1, California.)

Results were analyzed by regression analysis, using the IBM 701 computer at the University of California, Berkeley. Dependent variables were the dispersion ratio and the surface aggregation ratio. Independent variables used were the milliequivalents of the four cations Ca^{++} , Mg^{++} , K^+ , and Na^+ and various sums, squares, and cross-product functions of them. The estimating coefficients for each regression, and the Students *t* values associated with them are listed in Table 3.

Results. For data with so much variability,

TABLE 3. Regression Coefficients and t Values, Regression Constants, and Explained Variances in Equations Relating Erodibility Indices to Exchangeable Soil Cations

Equation	Dependent Variable	Constant Term	Regression Coefficients and <i>t</i> 's for Independent Variables										Explained Variance		
			Ca ⁺⁺	Mg ⁺⁺	K ⁺	Na ⁺	(Ca ⁺⁺ + Mg ⁺⁺)	(K ⁺ + Na ⁺)	(Ca ⁺⁺) ²	(Mg ⁺⁺) ²	(Ca ⁺⁺ + Mg ⁺⁺) ²	(K ⁺ + Na ⁺) ²		(Ca ⁺⁺) ²	(K ⁺ + Na ⁺) ²
(1)*	D.R.	† 108.7					-4.42 [2.98]	-45.29 [1.00]		+0.93 [2.44]	+38.58 [1.62]		-0.805 [0.55]		0.73
(3)	D.R.	80.1	-1.92 [2.73]	-1.17 [3.75]	+1.45 [0.15]	-93.64 [0.56]									0.68
(5)	D.R.	75.0					-1.44 [5.28]	-4.86 [0.60]							0.60
(7)	D.R.	90.7					-3.96 [3.00]		+0.68 [4.22]						0.67
(9)	D.R.	90.0	-3.78 [1.37]	-3.79 [2.29]				+0.74 [0.81]	+0.93 [1.73]			+0.43 [0.45]			0.73
(11)	D.R.	106.5	-3.57 [1.13]	-4.51 [2.43]				-47.48 [1.04]	+0.55 [0.41]	+106 [1.8]	+308 [1.06]	+0.93 [0.83]			0.75
(2)	S/A†	322.6					-20.28 [2.61]	-152.0 [0.64]			+450 [2.55]		-0.892 [0.12]		0.52
(4)	S/A	146.5	-6.26 [1.61]	-3.80 [2.22]	-20.14 [0.37]	+404.85 [0.44]									0.47
(6)	S/A	170.8					-4.09 [2.86]	-39.0 [0.92]							0.34
(8)	S/A	258.4					-19.21 [2.98]		+409 [2.40]						0.48
(10)	S/A	336.6	-34.15 [2.71]	-26.36 [3.54]				+1.217 [2.19]	+730 [2.98]				+0.652 [1.48]		0.62
(12)	S/A	371.7	-27.93 [1.93]	-29.04 [3.42]				-134.8 [0.65]	+993 [1.73]	+826 [3.05]	+57.55 [0.44]	+0.648 [1.26]			0.66

* Example; Equation 1.

Dispersion ratio = $108.7 - 4.42(\text{Ca}^{++} + \text{Mg}^{++}) - 45.29(\text{K}^+ + \text{Na}^+) + 0.093(\text{Ca}^{++} + \text{Mg}^{++})^2 + 38.58(\text{K}^+ + \text{Na}^+)^2 - 0.805(\text{K}^+ + \text{Na}^+)(\text{Ca}^{++} + \text{Mg}^{++})$.

† Dispersion ratio.

† Surface aggregation ratio.

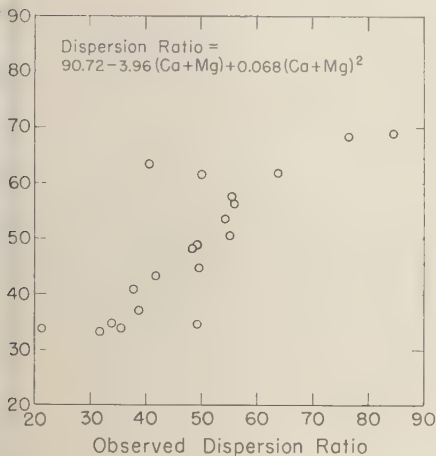


Fig. 1. Observed dispersion ratio plotted against calculated dispersion ratio.

observations is a minimum number for analysis. In spite of this difficulty, high correlation was found between cations and the soil erodibility indices. A study of the t values recorded in Table 3 shows, first, that the linear terms in exchangeable Ca^{++} , Mg^{++} , and Ca^{++} plus Mg^{++} are negatively correlated with the erodibility indices. Second, the curvilinear terms for Ca^{++} , Mg^{++} , and Ca^{++} plus Mg^{++} are significant in most models. Third, the small amounts of K^+ and Na^+ present in these samples have a statistically nonsignificant effect on the erodibility indices.

We checked the various models with a succession of F tests to determine whether the increased correlation of the larger models, that is, the addition of joint and separate relations, were significant. For an arbitrarily selected 5 per cent significance level, the best fit to the data was with an equation of the following type:

$$\text{Erodibility index} = a - b_1(\text{Ca}^{++} + \text{Mg}^{++}) + b_2(\text{Ca}^{++} + \text{Mg}^{++})^2$$

This equation is of the same type as equations 3 and 8 in Table 3.

In six regressions (equations 3, 9, 11, 4, 10, and 12) where Ca^{++} and Mg^{++} were separated, they were found to have similar coefficients and hence similar effects on the erosion indices.

A similar comparison between an equation of this type and a purely linear relationship showed

that the second-degree term should be included. Only one sample had a smaller dispersion ratio than the minimum value (33.1) developed by equation 7, and only two samples had smaller surface aggregation ratios than the minimum value (32.9) developed by equation 8.

The 29 per cent increase of explained variance (R^2) found when using equation 10 rather than equation 8 of Table 3 is statistically nonsignificant, although from the size of the t values it appears that all terms are important. A larger sample might show significance.

Discussion. The predictability of the dispersion and surface aggregation ratios from equations 7 and 8 is graphically shown in Figures 1 and 2. A rather high degree of association is indicated. Undoubtedly, much of the variability still remaining is associated with, and can be explained by, inherent differences in the geology of the parent material.

Possibly the observed correlation between erodibility and parent material could be accounted for by the fact that different parent materials form different kinds of clay minerals. Marshall [1949] has shown that cations may be sorbed by a clay with a wide range of bonding energies and that these bonding energies are functions of the type of colloid and of the pH. Further studies of soil erodibility should measure both of these independent variables as well as the amount and kind of exchangeable cations. From a complete study of these basic colloidal

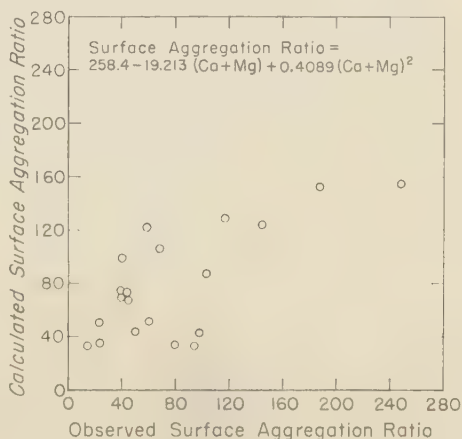


Fig. 2. Observed surface aggregation ratio plotted against calculated surface aggregation ratio.

properties it should be possible to develop chemical treatments that will change the erodibility of problem soils.

REFERENCES

- Anderson, H. W., Physical characteristics of soils related to erosion, *J. Soil and Water Cons.*, 6 (3), 129-133, July, 1951.
- Anderson, H. W., Suspended sediment discharge as related to streamflow, topography, soil, and land use, *Trans. Am. Geophys. Union*, 35 (2), 268-281, 1954.
- Jackson, M. L., *Soil Chemical Analysis*, Prentice-Hall, Englewood Cliffs, N. J., 500 pp., 1958.
- Marshall, C. E., *The Colloid Chemistry of the Silicate Minerals*, Academic Press, N. Y., 195 pp., 1947.
- Middleton, H. E., Properties of soils which influence soil erosion, *U. S. Dept. Agr. Tech. Bull.* 171, 1-16, 1930.
- Rost, C. O., and C. A. Rowles, A study of factors affecting the stability of soil aggregates, *Soil Sci. Am. Proc.*, 5, 421-433, 1941.
- Skula, K. P., and M. Nayar, Influence of Na, NH₄ and K ions on the permeability of calcium soil, *Current Sci. India*, 12 (5), 155-156, 1943.
- Smerdon, E. T., and R. P. Beasley, The tractive force theory applied to stability of open channels in cohesive soils, *Univ. Missouri Agr. Expt. St. Research Bull.* 715, 36 pp., 1959.

(Manuscript received October 31, 1960;
revised January 5, 1961.)

Heat Flow from a Differentiated Earth

SYDNEY P. CLARK, JR.

*Geophysical Laboratory, Carnegie Institution of Washington
Washington, D. C.*

Abstract. Calculations are made of the heat flow from an initially cold earth with radioactivity distributed uniformly through a surface shell, with account taken of radioactive decay. It is found that the heat flow exceeds the heat produced by about 10 per cent if the abundances of radioactive elements are the same as in chondrites and the radioactive shell is less than about 300 km thick. The heat flow calculated in this way exceeds the observed value, but neither figure is considered known accurately enough to warrant rejection of the chondrite model of the earth.

[Birch [1958] noted that if the earth has the same content of radioactive elements as the chondritic meteorites, the total rate of heat loss from the earth is roughly the same as its current rate of heat production, according to recent measurements of the radioactivity of meteorites. This model was taken to imply that the radioactive elements must be strongly concentrated toward the surface. MacDonald [1959] found, however, that if the radioactivity is distributed uniformly throughout the uppermost 600 km of the earth, the calculated heat flow exceeds the observed value (see his models 16, 17, and 18). The present work extends MacDonald's calculations to cases of more extreme upward concentration of radioactivity. The numerical results are in substantial agreement with MacDonald, but another way of interpreting the data is considered here.

There are two reasons why the loss of radioactive heat from the earth should differ from the present rate of heat production. Some of the heat currently being produced at depth causes heating and is thereby stored in the earth. On the other hand, radioactive heat production was greater in the past because of radioactive decay, and therefore the heat flow is greater than would be found if the production of heat were uniform in time. For convenience we shall refer to these opposing effects as the 'storage effect' and the 'decay effect,' respectively.

Since it is not obvious which effect is dominant in the earth, the question was investigated by examination of an earth model similar to those described by Jacobs [1956]. All the radioactive heat production is assumed to be uniformly distributed in a spherical shell extending from

the surface to some given depth. The initial temperature is assumed to be zero throughout, as is the surface temperature for all times. An explicit solution to this problem, in the form of a slowly convergent infinite series, is given by Jacobs (see Appendix). In the present study, heat flows at the surface were found by graphical differentiation of temperatures near the surface, which were calculated by an IBM-704 digital computer. This procedure is accurate enough for present purposes. A more rapidly convergent solution to this problem can be found, but it involves error functions of complex argument and is also very cumbersome.

The solution given by Jacobs [1956] was obtained on the assumption that both the thermal conductivity and thermal diffusivity are constant. This is plausible in the present case since we are mainly concerned with shallow depths and, consequently, relatively low temperatures. Under these conditions processes leading to strongly temperature-dependent thermal conductivity, such as radiative transfer, will not seriously affect the heat flow at the surface. On this model the heat flow depends explicitly on the thermal diffusivity, but not on the conductivity. The diffusivity was taken to be $36 \text{ km}^2/10^6 \text{ years}$, a value representative of rocks near the surface. The results are not sensitive to moderate changes in this quantity.

Other parameters that enter the problem are the decay constants and abundances of the important radioactive isotopes, the radius of the earth, the thickness of the radioactive shell, and the time since differentiation produced the shell. Decay constants and abundances of the radioactive isotopes are given in Table 1; the mean

TABLE 1. Decay Constants and Abundances of Radioactive Isotopes

Iso- tope	Decay Constant,* 10^{-10} years $^{-1}$	Present Abundance,† 10^{18} gm	Fraction of Present Heat Pro- duction
U ²³⁵	9.72	0.5	0.01
U ²³⁸	1.54	65.5	0.20
K ⁴⁰	5.305	500	0.58
Th ²³²	0.499	241	0.21

* Aldrich and Wetherill [1958].

† Birch [1958]. Based on an initial chondritic composition for the whole earth.

radius of the earth is 6371 km, and several values of thickness of the radioactive shell were assumed. The quantity tabulated below is the ratio of present heat flow to present heat production. Hence the absolute abundances of the radioactive isotopes do not enter the calculations, although the ratios of abundances do. Calculations were made for times of 2×10^9 , 3×10^9 , and 4×10^9 years. The oldest known rocks were formed about 3×10^9 years ago [Tilton and Davis, 1959], and the greatest age adopted approaches the potassium-argon ages of the stony meteorites [Geiss and Hess, 1958; Wasserburg and Hayden, 1955].

The results for the chondritic model, for K⁴⁰ alone, and for heat production independent of time are given in Table 2. The chondritic model after times of 3×10^9 years and 4×10^9 years is of greatest geophysical interest, since rocks considerably older than 2×10^9 years have been found, radioactive decay has certainly occurred, and uranium and thorium must contribute a substantial amount of radiogenic heat. The remaining entries in the table are interesting mainly for contrast; K⁴⁰ was singled out for special attention because it has the shortest half-life of the important heat-producing isotopes and therefore the decay effect is greatest for this nuclide.

Regarded as functions of thickness of the radioactive shell, the ratios in the last two columns of part A, Table 2, pass through broad maxima and appear to approach unity as the thickness of the shell approaches zero. The fact that in these cases the heat flow generally slightly exceeds the heat production, even if the radioactive layer is only 30 km thick, is the most

important result of the calculations. It implies that the decay effect more than compensates for the storage effect, and that the earth is, on the average, cooling.

If the whole earth has the same radioactive content as the chondritic meteorites, the heat flow in equilibrium with the heat production would be about 1.4×10^{-8} cal/cm² sec. If the mantle was initially chondritic in composition and the core initially contained no radioactivity, a figure about $\frac{2}{3}$ as large is obtained. According to the pertinent columns in Table 2, the storage and decay effects cause the heat flow to exceed its equilibrium value by about 10 per cent. Hence if all the radioactivity is presently within a few hundred kilometers of the surface, we should expect the radiogenic portion of the heat flow from an initially chondritic earth to be about 1.5×10^{-8} cal/cm² sec, and the heat flow from an earth with initially chondritic mantle and nonradioactive core to be about 1.0×10^{-8}

TABLE 2. Ratio of Heat Flow to Present Heat Production

Thickness of Shell, km	Time since Formation of Shell, years		
	2×10^9	3×10^9	4×10^9
A. Chondritic Model			
32	0.99	1.01	1.02
64	0.98	1.02	1.05
127	0.97	1.04	1.11
255	0.90	1.02	1.15
637	0.63	0.84	1.08
B. K ⁴⁰ Alone			
32	1.00	1.03	1.05
64	1.01	1.05	1.10
127	1.01	1.10	1.20
255	0.97	1.14	1.31
637	0.71	0.99	1.32
C. Heat Production Constant in Time			
32	0.97	0.97	0.98
64	0.93	0.94	0.95
127	0.87	0.90	0.92
255	0.76	0.80	0.84
637	0.48	0.56	0.61

*This is the time in the past at which the temperature is assumed to be zero everywhere. In the chondritic case the heat production is adjusted initially so that it equals the values given in Table 1 at the present time.

l/cm² sec. The average observed heat flow is commonly taken to be about 1.2×10^{-6} l/cm² sec.

No discrepancy between observed and computed heat flow arises in the case of an initially chondritic mantle. The calculations that lead to the results in Table 2 were made on the assumption that the temperature was zero at the time that upward concentration took place. This is not plausible in the actual earth, and heat flow due to initial heat must be added to the radioactive flux. This could bring the observed and computed heat flows into agreement.

In the case of an initially chondritic earth, a discrepancy between observed and computed heat flow arises unless the contribution from initial heat is very small. The average heat flow from the earth is uncertain by 20 per cent or so. The heat production in chondrites is uncertain by at least 10 per cent, and the number of determinations of uranium and thorium by neutron activation is still small. Because of these uncertainties, it cannot be concluded that the difference between the figures 1.5×10^{-6} cal/cm² sec and 1.2×10^{-6} cal/cm² sec is significant. Even if the calculated heat flow is increased to, say, 1.8×10^{-6} cal/cm² sec by initial heat, it is difficult to reconcile it with observations.

Of the assumptions made in calculating the heat flow, three seem especially tenuous, and these are the ones most critical to the results. These concern the validity of the chondrite model of the earth, the completeness with which radioactivity has been concentrated toward the surface, and the amount of heat flow to be expected from initial heat. If the earth contains on the average less radioactivity than the chondrites, or if some of this radioactivity remains at depths greater than about 600 km, the calculated heat flow can readily be brought into agreement with the observed value. These possibilities were both recognized by MacDonald [1959, p. 1995]. The discrepancy can be removed, or at worst kept small, if the contribution from initial heat is also small. MacDonald derived his temperatures at the time of upward concentration from a melting curve of diopside, extrapolated to high pressures. Temperatures obtained in this way are probably too high to represent the onset of melting in a primitive earth, which undoubtedly contained many phases; much lower temperatures, possibly not exceeding 1000°C in the

outermost 300 km of the earth, may have prevailed at the time of upward concentration.

APPENDIX

Let \bar{T} denote the Laplace transform of the temperature, T . We must solve

$$\frac{d^2(r\bar{T})}{dr^2} - q^2(r\bar{T}) = \frac{P_0 r}{K(p + \lambda)} \quad R > r > r_1$$

and

$$\frac{d^2(r\bar{T})}{dr^2} - q^2(r\bar{T}) = 0 \quad r_1 \geq r \geq 0$$

with \bar{T} and $d\bar{T}/dr$ continuous at $r = r_1$. Here R is the radius of the earth, and K , the thermal conductivity. The heat production per unit volume and unit time is $P_0 \exp(-\lambda t)$ in $R \geq r \geq r_1$ and zero in $r_1 \geq r \geq 0$, and q^2 stands for p , the parameter of the Laplace transform, divided by the thermal diffusivity.

The solution to these equations in the radioactive layer is

$$\begin{aligned} \bar{T} = \frac{P_0 \kappa}{K p (p + \lambda)} \left\{ 1 - (\sinh q r_1 \cosh q R \right. \\ \left. - q r_1 \cosh q r_1 \cosh q R + q R) \frac{\sinh q r}{q r \sinh q R} \right. \\ \left. + (\sinh q r_1 - q r_1 \cosh q r_1) \frac{\cosh q r}{q r} \right\} \end{aligned}$$

Application of the inversion theorem of the Laplace transform then leads to the equation [Jacobs, 1956]

$$T = \frac{2 P_0 R^3}{\pi K r} \sum_{n=1}^{\infty} (e^{-\lambda t} - e^{-\pi^2 n^2 \kappa t / R^2}) \sin \frac{n \pi r}{R} \cdot \left[\frac{(-1)^n + \frac{1}{n \pi} \sin \frac{n \pi r_1}{R} - \frac{r_1}{R} \cos \frac{n \pi r_1}{R}}{n(R^2 \lambda / \kappa - n^2 \pi^2)} \right]$$

A simpler expression for the heat flow at the surface can be found when $\lambda = 0$. At $r = R$,

$$\begin{aligned} K \frac{dT}{dr} = \frac{P_0}{R p q^2} \\ (\sinh q R - \sinh q r_1 \\ + q r_1 \cosh q r_1 - q R \cosh q R) \frac{1}{\sinh q R} \end{aligned}$$

A solution to this equation, useful for small values of the time, can readily be found by a

standard procedure [Carslaw and Jaeger, 1959]. If the hyperbolic functions are expressed in exponential form and the denominator expanded by the binomial theorem, the transform can be inverted by use of a table of Laplace transforms. The result is

$$\begin{aligned}
 K \frac{\partial T}{\partial r} = & \frac{P_0 \kappa t}{R} \\
 & \cdot \left\{ 1 - 4 \sum_{n=0}^{\infty} \left(i^2 \operatorname{erfc} \frac{(2n+1)R - r_1}{\sqrt{4\kappa t}} \right. \right. \\
 & \left. \left. - i^2 \operatorname{erfc} \frac{(2n+1)R + r_1}{\sqrt{4\kappa t}} \right) \right\} \\
 & + P_0 \sqrt{4\kappa t} \sum_{n=0}^{\infty} \left\{ \frac{r_1}{R} i \operatorname{erfc} \frac{(2n+1)R - r_1}{\sqrt{4\kappa t}} \right. \\
 & \left. + \frac{r_1}{R} i \operatorname{erfc} \frac{(2n+1)R + r_1}{\sqrt{4\kappa t}} \right. \\
 & \left. - i \operatorname{erfc} \frac{2nR}{\sqrt{4\kappa t}} - i \operatorname{erfc} \frac{(2n+2)R}{\sqrt{4\kappa t}} \right\}
 \end{aligned}$$

In cases of geophysical interest, terms with $n > 0$ need not be retained.

Acknowledgments. I am indebted to Francis Birch, Gordon J. F. MacDonald, and Eugene C. Robertson for helpful comments on this manuscript.

REFERENCES

- Aldrich, L. T., and G. W. Wetherill, Geochronology by radioactive decay, *Ann. Rev. Nuclear Sci.*, **8**, 257-298, 1958.
- Birch, F., Differentiation of the mantle, *Bull. Geo. Soc. Am.*, **69**, 483-485, 1958.
- Carslaw, H. S., and J. C. Jaeger, *Conduction of Heat in Solids*, 2d ed., Clarendon Press, Oxford, 1959.
- Geiss, J., and D. C. Hess, Argon-potassium age and the isotopic composition of argon from meteorites, *Astrophys. J.*, **127**, 224-236, 1958.
- Jacobs, J. A., The earth's interior, *Handbuch der Physik*, **47**, 364-406, 1956.
- MacDonald, G. J. F., Calculations on the thermal history of the earth, *J. Geophys. Research*, **64**, 1967-2000, 1959.
- Tilton, G. R., and G. L. Davis, *Geochronology, in Researches in Geochemistry*, John Wiley & Sons, New York, 1959.
- Wasserburg, G. J., and R. J. Hayden, Age of meteorites by the $A^{40}\text{-K}^{40}$ method, *Phys. Rev.*, **97**, 86-87, 1955.

(Manuscript received February 2, 1961.)

Paleomagnetic Study of the Sudbury Basin¹

PETER J. HOOD

Geophysics Laboratory, University of Toronto, Canada

Abstract. A reconnaissance paleomagnetic study has been carried out on the Precambrian nickel irruptive of the Sudbury basin. A total of 60 oriented samples was collected along a single traverse across the northern range and along four traverses across the southern range of the basin. The remanent magnetism (RM) vectors of the 266 specimen cubes cut from the samples were subsequently measured with a 'spinner' magnetometer. The RM directions for the south range are closely grouped on the Schmidt equal-area plot, and this fact indicates that there has been negligible relative rotary movement along the 16 miles of the south range sampled since the irruptive last cooled through the Curie point of its magnetic constituents. It was also found that the RM directions for the N and S ranges differ by approximately 40°, and this may be explained by assuming that there has been 40° of relative tilting about an E-W axis through the middle of the basin. In three of the S range traverses the susceptibility (κ) of the collected samples was higher on the S end of the traverse than on the N end, and in general the κ 's obtained for the N range were much higher than those obtained on the S range. It would appear that the magnetic anomalies observed over the norite sections of the N and S ranges are mainly due to the presence of RM and to a lesser extent to an induced magnetization. The pole position calculated from the RM directions is dependent on the amount of assumed relative tilting. However if the rotation is equally divided between the two ranges, the resultant pole position is 38.4°N, 99.4°W, which is in the central U. S.

Introduction—general geology. The Sudbury area is one part of a geologic province extending more than 600 miles from the west end of Lake Superior, easterly through Sudbury and Cobalt, and on into Quebec. This vast belt is characterized by the occurrence of a quartz-diorite type of rock with which are associated important metallic ores.

Structural studies have shown that the nickel irruptive of the Sudbury basin, with which the nickel-copper deposits are associated, has the shape of an asymmetric basin (Fig. 1), its rim being exposed at the surface in the form of an elliptical ring elongated in a northeastern direction. The long and short axes of this ring are approximately 37 and 17 miles long, respectively. The results from a gravity survey of the Sudbury basin [Miller and Innes, 1955] have appeared to support the contention by Collins and Kindle [1935] that the structure is approximately an asymmetric syncline.

Age determinations by the potassium-argon and rubidium-strontium methods [Wetherill, Davis, and Aldrich, 1957] have indicated that the nickel irruptive was emplaced 1200 to 1800 million years ago.

The apparent thickness of the nickel irruptive, estimated from outcrops and from near-surface dips, ranges from 1 to 2 miles. It consists of an inner layer of micropegmatite separated by a transition zone from an outer, somewhat narrower layer of norite. Along most of the northern side of the basin the base of the norite near the surface dips inward at approximately 35° [Thomson, 1956], while at several places along the eastern and southeastern sides of the nickel irruptive the outer unfaulted contact is extremely steep, dipping outward in several places [Lochhead, 1955].

Collins [1934] described the norite as being a massive, dark grey rock, as coarse as ordinary granite, consisting of orthorhombic and monoclinic pyroxene, hornblende, biotite, plagioclase (basic labradorite to andesine), quartz, titaniferous magnetite, yellow sulfide (pyrite and pyrrhotite according to Phemister [1925]), titanite, and apatite. Collins found the micropegmatite to be more uniform than the norite. It is usually a dull, salmon-pink rock, somewhat coarser than the norite, composed chiefly of sodic plagioclase, orthoclase or microcline, quartz, and biotite.

The space inside the basin structure is occupied by a series of volcanic and sedimentary rocks, the layering of which outcrops in concentric

¹ Based on a thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at the University of Toronto.



Fig. 1. Geologic map of the Sudbury basin showing sample traverses (After map 1956-1, Ontario Department of Mines).

rings. These volcanic and sedimentary rocks are considered to be conformable with the nickel irruptive.

North of the intrusive is an extensive area of granite, granite gneiss, and highly granitized sediments. To the south is a succession of volcanic, metamorphic, and sedimentary rocks, highly folded and faulted, striking in general east-northeast and dipping at steep angles mainly to the north.

Although the Sudbury area has been the subject of long and intensive studies by both mining companies and government organizations, the geology of the basin is still one of vigorous controversy. This study was carried out in the hope of throwing more light on the tectonic history of the area.

Sampling and measurement. A total of 60 oriented samples was collected for paleomagnetic studies from the nickel irruptive of the Sudbury basin during 1957; 37 of the samples were from the norite part of irruptive (Fig. 1). All the samples were taken either from road-cuts or from outcrops close to the highway. A suite of 18 samples was collected on the north range near the village of Levack, and the rest of the

samples were collected on the south range at Creighton, Azilda, Blezard, and Garson.

In general, a sample sufficiently large to yield about five 1-inch cubes was collected, and the samples were oriented with a Brunton compass, care being taken to see that it was not appreciably deflected by the rock in the outcrop.

A total of 266 specimen cubes was cut with diamond saws from 59 of the samples. The natural remanent magnetization of each specimen cube was then determined in a 'spinner' magnetometer [Hood, 1958]. This apparatus was of the conventional 'rock generator' type in which the rock specimen is rotated opposite a fixed coil system at 45 revolutions per second, compressed air being used as a source of power.

Results. The directions of remanent magnetization of the specimens obtained from the five traverses have been plotted on the Schmidt equal-area projection [Graham, 1949] in Figure 2.

From the Azilda projection, it can be seen that the directions for the majority of the cubes fall in a fairly close group; however 11 specimens all from the same road cut, do not fall in the group, and it is reasonable to suppose that they may have been some relative movement between

and the main mass of norite. The results from his outcrop, therefore, have not been used in the calculation of the mean direction. Five of the cubes were demagnetized in an a-c field of about 125 oersteds using *Thellier and Rimbart's* [1954] method. The specimens, with one exception, were only slightly affected, which indicated that they possessed good stability and that the magnetization was of the thermoremanent type. The exception, which initially lay a little outside the group, had its direction considerably improved by a-c demagnetization, and was brought into the group. Four of the specimens from the anomalous outcrop previously mentioned, were also little affected by a-c demagnetization.

The Schmidt projections for the four traverses, Creighton, Blezard, Garson, and Levack, show fairly good groupings of the results. A number of the specimens from these traverses were also demagnetized in order to improve the grouping.

To determine the variation in susceptibility (κ) across the nickel irruptive, 'chip samples' were obtained from a number of the samples collected. The chips from a given sample were then crushed and divided into portions. The κ of each portion was measured with a calibrated susceptibility meter of the alternating current type, and the results were averaged to obtain the mean κ of the collected sample.

Figure 3 shows the results of the susceptibility measurements together with the average values of the intensity of remanent magnetization (RM) of the samples. Q is the Koenigsberger ratio, the ratio of remanent to induced magnetization; the value of the earth's field at Sudbury was taken to be 0.59 oersteds. Results from the Garson traverse were not included in Figure 3 because of the paucity of samples collected.

For the Azilda, Blezard, and Creighton traverses, it can be seen that all three physical properties are rather variable within the exposed section of the norite. In general, all these quantities decrease rapidly as the transition zone is approached. For all three traverses, κ tends to decrease from south to north across the norite, and for the Creighton and Blezard samples there appear to be well-defined maxima in Q and RM toward the center of the norite. Q remains well above the value of unity along all of the traverses and reaches values of between 10 and 20 within the norite.

Discussion of results and conclusions. The

mean value of κ of the transition zone at Levack is about a factor of 100 higher than was obtained for the transition zone on the south range. However, the κ of the samples from the norite is similar to that obtained on the south range. Examination of an aeromagnetic map covering the area south from Levack to the south range [Zietz and Henderson, 1955] also shows that the magnetic anomalies produced by the transition zone on the north range have a much greater amplitude than those associated with the transition zone on the south range.

In addition, Q is less than unity for the transition zone near Levack. It would appear therefore, that the transition zone at Levack contains a greater percentage of magnetite than the transition zone on the south range, since κ is usually directly dependent on the magnetite content [Mooney and Bleifuss, 1953; Runcorn, 1956]. Thus the transition zone at Levack also contains a rather higher percentage of magnetite than the norite and micropegmatite from which it is thought to be formed. This is consistent with the results of Collins [1934], who showed that at the junction of the norite and the transition zone at Levack there is a concentration of titaniferous magnetite.

Comparison of the results from Azilda, Blezard, and Creighton (Fig. 3) shows that κ is higher on the extreme southern edge of the norite, in each case, than on the northern edge. This would, of course, be consistent with some differentiation of the magnetite across the norite. The rise in the value of κ toward the center could be attributed to either an increase in the percentage composition or in the grain size of the ferromagnetic constituents.

Examination of the Schmidt projections shows that the directions of the RM vectors of the south range specimens are well grouped (see also Fig. 4) and do not differ by more than 40° from the present direction of the earth's field. As a result, the component of the mean RM vector in the direction of the earth's magnetic field will be considerable. The magnetic anomalies observed over the norite section of the nickel irruptive on the south range are therefore due largely to the presence of RM in the rocks, and to a lesser extent to induced magnetization because the value of Q remains well above unity across the three traverses. This statement also appears to be true for the norite on the north

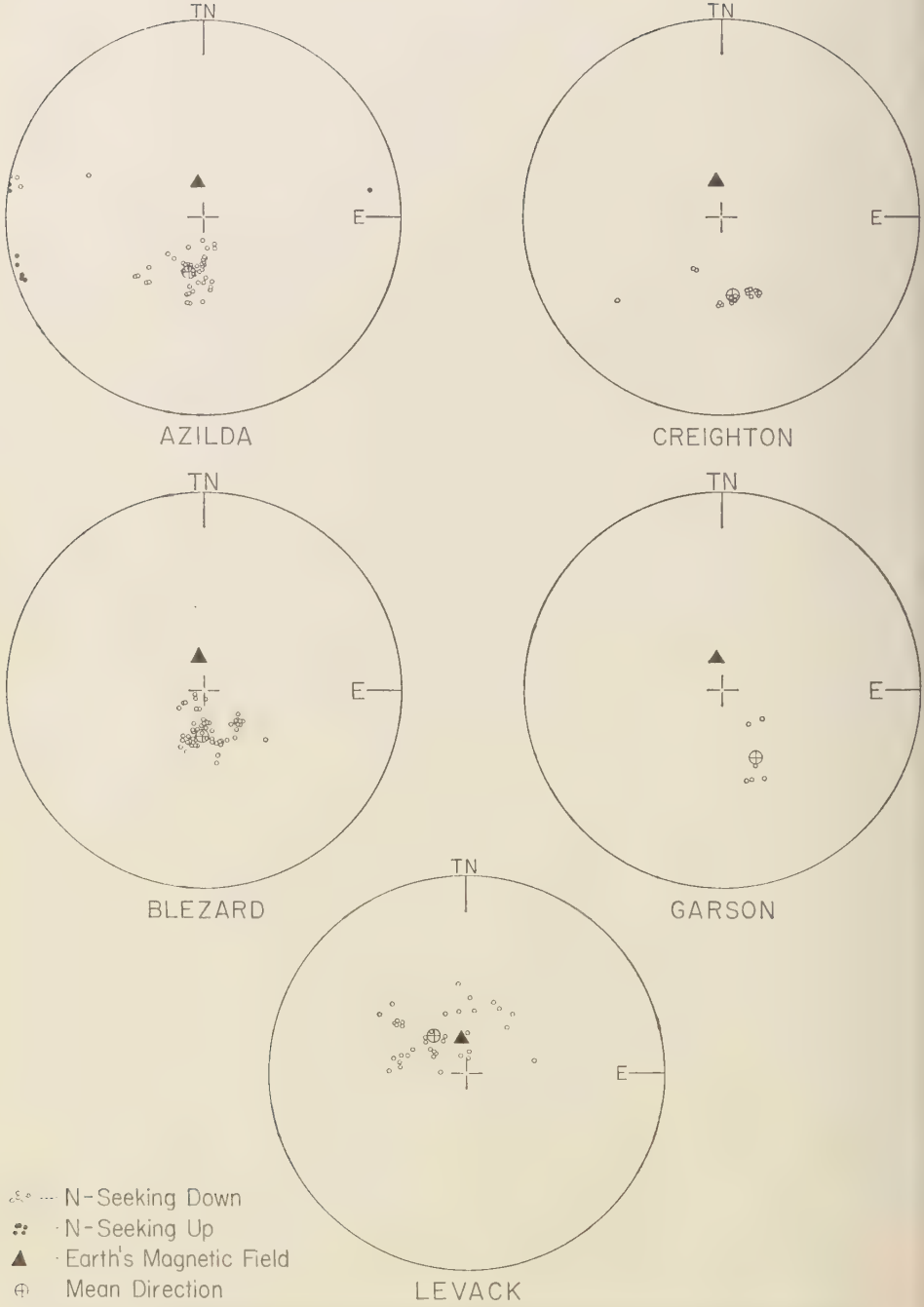


Fig. 2. Schmidt equal-area projections for the five traverses across the nickel irruptive.

range, i.e., Levack, although this is probably not the case for the transition zone or the micropegmatite, since Q is rather less than unity for these.

Thus these results provide a further example that RM is rather more important than is generally realized by those engaged in the interpretation of aeromagnetic maps. Several articles [Sutton and Mumme, 1957; Hall, 1959; Green, 1960; Girdler and Peter, 1960] have recently been published which have also shown that it is often unjustified to assume that the effects of RM can be ignored in the interpretation of aeromagnetic surveys.

Figure 4 is a Schmidt equal-area plot of the combined RM results for the south range. Since

the grouping is quite good, and since the mean directions for each of the four traverses fall close to one another, it would appear that there has been very little relative rotary movement across the 16 miles of the nickel irruptive between Creighton and Garson since the irruptive last cooled through the Curie point of its magnetic constituents.

A summary of the paleomagnetic results is given in Table 1, together with the results of the statistical analysis [Fisher, 1953] which is usually employed in paleomagnetic investigations.

The Precambrian geomagnetic pole positions have been calculated from the mean directions of the results from Levack and the south range [Irving, 1956], and these are also listed in Table 1.

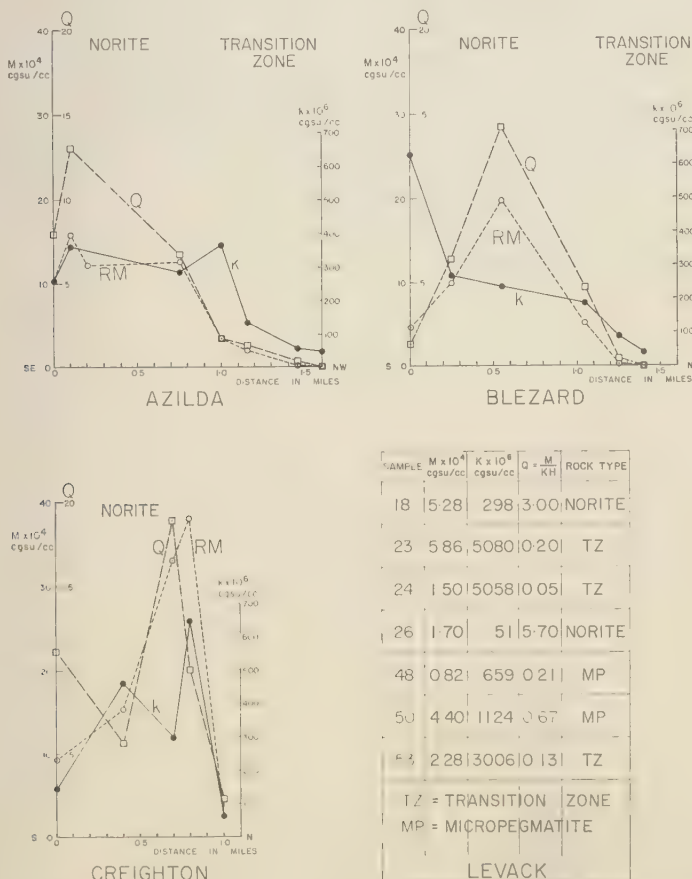


Fig. 3. Susceptibility (κ), remanent magnetization (RM), and the Koenigsberger ratio (Q) for four traverses.

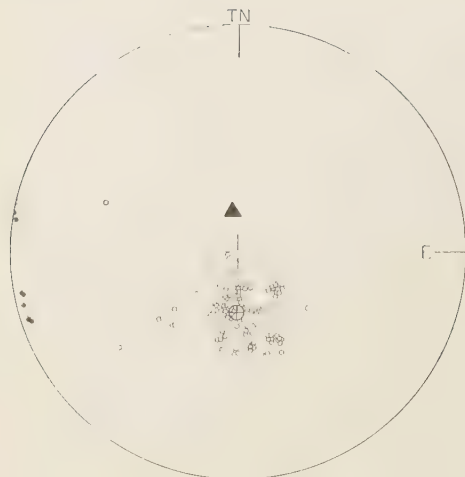


Fig. 4. Schmidt projection for the south range.

The most important result of this investigation is that the mean directions of the samples collected from the north and south ranges differ by some 40° ; the mean direction for the north range (Levack) dips to the north, and that for the south range dips to the south.

The geologic interpretation of this result is

equivocal. It seems reasonable to assume that the shearing on the south range has not modified the RM directions obtained from the four traverses which extend over a distance of 10 miles. Therefore, if the eruptives on either side of the basin are the same geologic age, it would appear that there has been some relative movement about an east-west axis through the center of the basin, between the eruptive on the north and the eruptive on the south range. In fact there appears to have been about 40° of relative tilting.

In calculating a pole position for the combined Sudbury results, if it is assumed that there has been equal tilting, i.e. 20° , on either side of the axis, the result labeled (a) in Table 1 is obtained. Alternatively, there is some indication from geologic evidence [Thomson, 1956] that the south range may have been more affected by tectonic action than the north range, since the latter acted as a foreland. If we therefore assume that the north range was only tilted by 10° and the south range by about 30° , the result (b) in Table 1 is obtained. This pole position, however, falls farther away from those obtained by other workers [e.g. Graham, 1953; Irving, 1957] than does result (a).

All these conclusions are based on the assumption

TABLE 1. Paleomagnetic Results from Sudbury

Sample Location	Magnetic Direction			R^\dagger	α_{95}^\ddagger	K^\S	Sampling Area		Pole position			
	Decl.	Incl.	N^*				Long., °W	Lat., °N	Long., °W	Lat., °N	Oval of Confidence	
Levack	320.4	69.9	41	38.980	5.2	19.8	81.4	46.6	140.7	64.0	8	9
Creighton	171.4	57.0	26	25.464	4.2	46.7	81.2	46.5
Azilda	194.4	66.7	45	45.342	2.6	66.8	81.1	46.5
Blezard	183.9	71.7	78	77.105	1.7	86.0	81.0	46.6
Garson	155.7	59.2	6	5.887	10.2	44.2	80.9	46.6
South range combined	182.8	68.0	155	151.865	1.6	49.1	81.0	46.5	82.8	7.6	2	3
North and south ranges combined												
(a)	245.3	82.3	196	81.2	46.6	99.4	38.4
(b)	300	78	196	81.2	46.6	115.9	53.4

* Number of specimen cubes measured.

† Resultant vector of N unit vectors.

‡ Angular radius of the circle of confidence for a probability of 95 per cent.

§ Dispersion.

|| Δp , semi-major axis; Δm , semi-minor axis.

(a) Assuming equal tilting of north and south ranges.

(b) Assuming 10° tilt of north range and 30° tilt of south range.

on that the direction of RM observed actually coincided with that of the earth's field at the time the rocks cooled through their Curie points. It is not inconceivable that some other process, as yet undetermined, has produced this difference in the direction of the RM vectors in the north and south ranges. For this and other reasons, it is felt that this paleomagnetic study of the nickel irruptive at Sudbury has really been only a reconnaissance study.

For this study samples from rock outcrops only were collected. Additional information on the tectonic history of the basin might be forthcoming if the intrusive were also sampled at depth and at points closer to the more inaccessible southwest and northeast ends of the basin, and a more mineralogical investigation of the samples were undertaken.

I should like to thank Professors J. T. Wilson and R. M. Farquhar for much encouragement and helpful discussion, the Shell Oil Company limited for a fellowship during the year 1956-1957, and Hunting Survey Corporation Limited for the loan of a susceptibility meter.

REFERENCES

Collins, W. H., Life-history of the Sudbury nickel irruptive, I, Petrogenesis, *Trans. Roy. Soc. Can.*, IV, 28, 123-178, 1934.
 Collins, W. H. and E. D. Kindle, The life history of the Sudbury nickel irruptive II, *Trans. Roy. Soc. Can.*, IV, 29, 27-38, 1935.
 Fisher, R. A., Dispersion on a sphere, *Proc. Roy. Soc., London, A*, 217, 295-305, 1953.
 Girdler, R. W., and G. Peter, An example of the importance of natural remanent magnetization in the interpretation of magnetic anomalies, *Geophys. Prospecting*, 8, 474-483, 1960.
 Graham, J. W., The stability and significance of magnetism in sedimentary rocks, *J. Geophys. Research*, 54, 131-167, 1949.
 Graham, J. W., Changes of ferromagnetic minerals

and their bearings on magnetic properties of rocks, *J. Geophys. Research*, 58, 243-260, 1953.
 Green, R., Remanent magnetization and the interpretation of magnetic anomalies, *Geophys. Prospecting*, 8, 98-110, 1960.
 Hall, D. H., Direction of polarization determined from magnetic anomalies, *J. Geophys. Research*, 64, 1945-1959, 1959.
 Hood, P. J., Palaeomagnetic studies of some Precambrian rocks in Ontario, Unpublished Ph.D. thesis, University of Toronto, 1958.
 Irving, E., Palaeomagnetic and palaeoclimatological aspects of polar wandering, *Geofis. pura e appl.*, 33, 23-41, 1956.
 Irving, E., The origin of the palaeomagnetism of the Torridonian sandstones of N.W. Scotland, *Phil. Trans. Roy. Soc. London, A*, 250, 100-110, 1957.
 Lochhead, D. R., The Falconbridge ore deposit, Canada, *Econ. Geol.*, 50, 42-50, 1955.
 Miller, A. H., and M. J. S. Innes, Gravity in the Sudbury basin and vicinity, *Publ. Dominion Observatory*, 18, 13-43, 1955.
 Mooney, H. M., and R. Bleifuss, Magnetic susceptibility measurements in Minnesota, 2, Analysis of field results, *Geophysics*, 18, 383-393, 1953.
 Phemister, T. C., Igneous rocks of Sudbury and their relation to the ore deposits, *Ann. Rept. Ontario Dept. Mines*, 34, 1-57, 1925.
 Runcorn, S. K., Magnetization of rocks, in *Handbuch der Physik*, Springer-Verlag, Berlin, 47, 470-497, 1956.
 Sutton, D. J., and W. G. Mumme, The effect of remanent magnetization on aeromagnetic interpretation, *Australian J. Phys.*, 10, 547-557, 1957.
 Thellier, E., and F. Rimbart, Sur l'analyse d'aiguillettes fossiles par action de champs magnetiques alternatifs, *Compt. rend.*, 239, 1399-1401, 1954.
 Thomson, J. E., Geology of the Sudbury basin, *Ann. Rept. Ontario Dept. Mines*, 65, 1-56, 1956.
 Wetherill, G. W., G. L. Davis, and L. T. Aldrich, Age measurements on rocks north of Lake Huron, *Trans. Am. Geophys. Union*, 38, 412, 1957.
 Zietz, I., and R. G. Henderson, The Sudbury aeromagnetic map as a test of interpretation methods, *Geophysics*, 20, 307-317, 1955.

(Manuscript received October 28, 1960;
 revised January 10, 1961.)

Earth Currents of Deep Internal Origin

P. H. ROBERTS AND F. J. LOWES

*Department of Physics, King's College
Newcastle upon Tyne, England*

Abstract. The current system flowing in the earth's core that is responsible for the dipole geomagnetic field is toroidal and does not appreciably extend into the mantle. However, the mechanism supporting these currents almost certainly generates in addition a poloidal current system that does extend into the mantle and that could, in principle, be measured at the earth's surface. In this paper the current distribution in the mantle is examined for several assumed distributions of conductivity in the mantle. The effect of the oceans is also briefly considered. It is found that the potential gradient at the surface of the earth associated with these deep earth currents may be of the order of a millivolt/kilometer if thermoelectric emf's at the core boundary contribute appreciably to the magnetic field, but will be smaller if the primary emf's are inside the core.

1. **Introduction.** For many years measurements have been made of the potential gradients associated with the electric currents flowing in the crustal rocks and water near the earth's surface. These 'earth currents' originate mainly from sources near to or above the earth's surface: e.g., currents induced by the motion of water through the geomagnetic field; currents induced in the mantle by currents in or above the ionosphere; currents generated electrochemically by natural and man-made bodies in the crustal rocks; leakage currents from transmission lines, electric railways, etc.

In this paper the possible existence of steady electric currents flowing near the earth's surface, originating from the earth's core and associated with the mechanism maintaining the main geomagnetic field is considered. We ignore all sources of emf within the mantle; in particular we ignore any emf induced by varying fields. When steady currents must be poloidal and must therefore be associated with toroidal magnetic fields, i.e. fields which have no radial component anywhere. Such toroidal fields probably exist in the lower mantle and may possibly be of the order of tens of gauss, but they must fall to zero on the earth's surface and are almost certainly too weak to be detected near the earth's surface. However, the associated electric currents may be observable. If so, it should be possible to deduce, from a survey of the electric

current system at the surface of the earth, the general structure of the toroidal magnetic fields deep within the earth. Also, if the variation of electrical conductivity within the mantle were known sufficiently well, the magnitude, as well as the direction, of the electric and magnetic fields at the base of the mantle could be deduced, leading to an increased understanding of the mechanism supporting the main geomagnetic field.

To investigate these ideas quantitatively, we shall assume that the electrical conductivity σ of the mantle varies only with distance r from the geocenter. The spherical harmonics of the fields can then be considered separately. For a particular harmonic S_n , in any arbitrary system of spherical polar coordinates, we find (see eqs. 5, 7, 11 below) that the tangential components of the electric field \mathbf{E} are

$$E_\theta = E_n(r) \partial S_n / \partial \theta, \quad (1)$$

$$E_\phi = [E_n(r) / \sin \theta] \partial S_n / \partial \phi$$

and the components of the magnetic field \mathbf{H} are

$$H_r = 0,$$

$$H_\theta = [H_n(r) / \sin \theta] \partial S_n / \partial \phi, \quad (2)$$

$$H_\phi = -H_n(r) \partial S_n / \partial \theta$$

The results of our mathematical discussions will be summarized by the nondimensional ratios

$$f_n = \frac{\left| \frac{E_n(b)}{E_n(a)} \right|}{\frac{\text{Magnitude of tangential component of electric field at bottom of mantle}}{\text{Magnitude of tangential component of electric field at top of mantle}}} \quad (3)$$

$$N_n = \frac{1}{4\pi b\sigma(b)} \left| \frac{H_n(b)}{E_n(a)} \right| = \frac{1}{4\pi b\sigma(b)}$$

$$\times \frac{\text{Magnitude of tangential magnetic field at bottom of mantle}}{\text{Magnitude of tangential component of electric field at top of mantle}} \quad (4)$$

both for the harmonic of order n . (Electromagnetic units are used throughout, and the permeability is taken to be unity. a and b are the radius of the mantle and the core, respectively.)

2. *Mathematical theory: general solutions.* We assume that there is no source of emf in the mantle. The steady currents flowing in the mantle must then be composed of leakage currents from the core (and, possibly, currents generated thermoelectrically at the core-mantle interface). Thus, as we remarked earlier, the electric field \mathbf{E} and electric current density \mathbf{j} must be poloidal, and the magnetic field \mathbf{H} must be toroidal. The fields can be related to a function $T(r, \theta, \phi)$ by

$$\begin{aligned} \mathbf{H} &= H_0 \text{curl} [\mathbf{r}T(r, \theta, \phi)] = H_0 \text{grad } T \mathbf{r} \\ &= H_0(0, \partial T/\sin \theta \partial \phi, -\partial T/\partial \theta) \end{aligned} \quad (5)$$

$$\begin{aligned} \mathbf{j} &= (H_0/4\pi) \text{curl}^2 [\mathbf{r}T(r, \theta, \phi)] \\ &= (H_0/4\pi) [\text{grad } \partial(rT)/\partial r - \mathbf{r} \nabla^2(rT)] \\ &= \frac{H_0}{4\pi} \left[-\frac{1}{r \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial T}{\partial \theta} \right) \right. \\ &\quad \left. - \frac{1}{r \sin^2 \theta} \frac{\partial^2 T}{\partial \phi^2}, \frac{1}{r} \frac{\partial}{\partial \theta} \left(\frac{\partial}{\partial r} (rT) \right), \right. \\ &\quad \left. \frac{1}{r \sin \theta} \frac{\partial}{\partial \phi} \left(\frac{\partial}{\partial r} (rT) \right) \right] \end{aligned} \quad (6)$$

$$\begin{aligned} \mathbf{E} &= (H_0/4\pi\sigma) \text{curl}^2 [\mathbf{r}T(r, \theta, \phi)] \\ &= (H_0/4\pi\sigma) [\text{grad } \partial(rT)/\partial r - \mathbf{r} \nabla^2(rT)] \\ &= \frac{H_0}{4\pi\sigma} \left[-\frac{1}{r \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial T}{\partial \theta} \right) \right. \\ &\quad \left. - \frac{1}{r \sin^2 \theta} \frac{\partial^2 T}{\partial \phi^2}, \frac{1}{r} \frac{\partial}{\partial \theta} \left(\frac{\partial}{\partial r} (rT) \right), \right. \\ &\quad \left. \frac{1}{r \sin \theta} \frac{\partial}{\partial \phi} \left(\frac{\partial}{\partial r} (rT) \right) \right] \end{aligned} \quad (7)$$

Also, since the electric field is steady, it must be expressible as the gradient of a single-valued electrostatic potential $\psi(r, \theta, \phi)$:

$$\begin{aligned} \mathbf{E} &= -\nabla_0 \text{grad } \psi(r, \theta, \phi) = V_0(-\partial \psi/\partial r, \\ &\quad -\partial \psi/r \partial \theta, -\partial \psi/r \sin \theta \partial \phi) \end{aligned} \quad (8)$$

Here the functions T and ψ are nondimensional, the constants H_0 and V_0 having dimensions of magnetic field and electric potential, respectively.

On the assumption that σ depends only on r , the equivalence of (7) and (8) for \mathbf{E} requires that

$$V_0 \psi = -(H_0/4\pi\sigma) \partial(rT)/\partial r$$

and hence that

$$\begin{aligned} \sigma \frac{\partial}{\partial r} \left(\frac{1}{r} \frac{\partial}{\partial r} (rT) \right) \\ + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} (rT) \right) \\ + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2}{\partial \phi^2} (rT) = 0 \end{aligned} \quad (9)$$

or

$$\begin{aligned} \frac{1}{r^2 \sigma} \frac{\partial}{\partial r} \left(r^2 \sigma \frac{\partial \psi}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial \psi}{\partial \theta} \right) \\ + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2 \psi}{\partial \phi^2} = 0 \end{aligned} \quad (10)$$

If T and ψ are expanded in spherical harmonics

$$T(r, \theta, \phi) = \sum_{n=1}^{\infty} R_n(r) S_n(\theta, \phi) \quad (11)$$

$$\psi(r, \theta, \phi) = \sum_{n=1}^{\infty} \Psi_n(r) S_n(\theta, \phi) \quad (12)$$

Then, since the spherical harmonics S_n form a complete orthogonal set over the surface of a sphere, and satisfy

$$\begin{aligned} \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial S_n}{\partial \theta} \right) \\ + \frac{1}{\sin^2 \theta} \frac{\partial^2 S_n}{\partial \phi^2} = -n(n+1) S_n \end{aligned} \quad (13)$$

equations 9 and 10 require that

$$\frac{d}{dr} \left(\frac{1}{\sigma} \frac{d}{dr} (rR_n) \right) - \frac{n(n+1)}{r^2} (rR_n) = 0 \quad (14)$$

$$\frac{1}{r^2 \sigma} \frac{d}{dr} \left(r^2 \sigma \frac{d\Psi_n}{dr} \right) - \frac{n(n+1)}{r^2} \Psi_n = 0 \quad (15)$$

for each n .

When the variation of σ with r is given, (14) and (15) become equations defining the permissible radial potential functions R_n and Ψ_n .

It is interesting to notice that if R_{n1} and R_{n2} are solutions of (14), and Ψ_{n1} and Ψ_{n2} are solutions of (15) then

$$\begin{aligned} R_{n1}(r) \frac{d}{dr} [rR_{n2}(r)] \\ - rR_{n2}(r) \frac{d}{dr} [rR_{n1}(r)] = X_n \sigma(r) \end{aligned} \quad (16)$$

$$R_{n1}(r) \frac{d\Psi_{n2}(r)}{dr} - \Psi_{n2}(r) \frac{d\Psi_{n1}(r)}{dr} = \frac{Y_n}{r^2 \sigma(r)} \quad (17)$$

where X_n and Y_n are constants depending upon the choice of R_{n1} , R_{n2} and Ψ_{n1} , Ψ_{n2} and are nonzero if these are independent solutions of (14) and (15).

The particular solution of equation 14 which we require is one for which the normal flow of current at the earth's surface, $j_r(a, \theta, \phi)$, is zero, so that $R_n(a) = 0$. Such a solution can be expressed in terms of any two independent solutions R_{n1} and R_{n2} as

$$R_n(r) = C_n [R_{n1}(a)R_{n2}(r) - R_{n2}(a)R_{n1}(r)] \quad (18)$$

and this implies

$$\begin{aligned} \Psi_n(r) = - \left(\frac{H_0}{V_0 \sigma(r)} \right) \frac{C_n}{4\pi} \left[R_{n1}(a) \frac{d}{dr} (rR_{n2}(r)) \right. \\ \left. - R_{n2}(a) \frac{d}{dr} (rR_{n1}(r)) \right] \end{aligned} \quad (19)$$

Expressing E_θ , E_ϕ , H_θ , H_ϕ in the forms (1), (2), we now find

$$M_n = \frac{|a\Psi_n(b)|}{|b\Psi_n(a)|} = \frac{a\sigma(a)}{b\sigma(b)} \left| \frac{R_{n1}(a) \left[\frac{d}{dr} (rR_{n2}(r)) \right]_b - R_{n2}(a) \left[\frac{d}{dr} (rR_{n1}(r)) \right]_b}{R_{n1}(a) \left[\frac{d}{dr} (rR_{n2}(r)) \right]_a - R_{n2}(a) \left[\frac{d}{dr} (rR_{n1}(r)) \right]_a} \right| \quad (20)$$

or, alternatively, using (16),

$$M_n = \left(\frac{a}{b} \right)^2 \left| \frac{R_{n1}(a) \left[\frac{d}{dr} (rR_{n2}(r)) \right]_b - R_{n2}(a) \left[\frac{d}{dr} (rR_{n1}(r)) \right]_b}{R_{n1}(b) \left[\frac{d}{dr} (rR_{n2}(r)) \right]_b - R_{n2}(b) \left[\frac{d}{dr} (rR_{n1}(r)) \right]_b} \right| \quad (21)$$

where the subscript a on $[\dots]_a$ indicates that the quantity between the brackets is to be evaluated at $r = a$ (and similarly for the subscript b).

In a similar way we find

$$N_n = \frac{a\sigma(a)}{b\sigma(b)} \left| \frac{R_{n1}(a)R_{n2}(b) - R_{n2}(a)R_{n1}(b)}{R_{n1}(a) \left[\frac{d}{dr} (rR_{n2}(r)) \right]_a - R_{n2}(a) \left[\frac{d}{dr} (rR_{n1}(r)) \right]_a} \right| \quad (22)$$

$$N_n = \left(\frac{a}{b} \right)^2 \left| \frac{R_{n1}(a)R_{n2}(b) - R_{n2}(a)R_{n1}(b)}{R_{n1}(b) \left[\frac{d}{dr} (rR_{n2}(r)) \right]_b - R_{n2}(b) \left[\frac{d}{dr} (rR_{n1}(r)) \right]_b} \right| \quad (23)$$

3. *Mathematical theory: effect of sea.* We now discuss the effects of the oceans. These will be represented by a thin shell of thickness βa and of constant conductivity S covering the surface of the earth. For mathematical convenience we will then take the limit $\beta \rightarrow 0$, $S \rightarrow \infty$ in such a way that

$\beta S \rightarrow \mu \sigma(a)$, where $\sigma(a) = \lim_{r \rightarrow a-0} \sigma(r)$ is the conductivity of the mantle immediately below the ocean [cf. *Lahiri and Price*, 1939]. Subject always to the proviso that $\sigma(r)$ does not vanish anywhere within the mantle, we have the following scheme of solutions:

In the oceans, $a \leq r \leq a(1 + \beta)$,

$$\left. \begin{aligned} R_n(r) &= C \{ [r/a(1 + \beta)]^n \\ &\quad - [a(1 + \beta)/r]^{n+1} \} \\ \Psi_n(r) &= -(H_0/V_0 S)(C/4\pi r) \\ &\quad \cdot \{ (n+1) [r/a(1 + \beta)]^n \\ &\quad + n[a(1 + \beta)/r]^{n+1} \} \end{aligned} \right\} \quad (24)$$

In the mantle, $b \leq r \leq a$,

$$\left. \begin{aligned} R_n(r) &= A_1 R_{n1}(r) - A_2 R_{n2}(r) \\ \Psi_n(r) &= -\frac{H_0}{V_0 \sigma(r)} \cdot \frac{1}{4\pi r} \left[A_1 \frac{d}{dr} (r R_{n1}(r)) \right. \\ &\quad \left. - A_2 \frac{d}{dr} (r R_{n2}(r)) \right] \end{aligned} \right\} \quad (25)$$

where A_1, A_2 are constants to be determined from the constant C by the requirements that both $R_n(r)$ and $\Psi_n(r)$ are continuous at $r = a$. These conditions automatically ensure that $\mathbf{j} \cdot \mathbf{r}$ is continuous at $r = a$. The solution (24) satisfies the requirement that $\mathbf{j} \cdot \mathbf{r}$ be zero on $r = a(1 + \beta)$. Applying the boundary conditions at $r = a$ we find to first order in β

$$A_1 = -C(2n+1) \frac{\sigma(a)}{S} \cdot \frac{R_{n2}(a) + \mu \left[\frac{d}{dr} (r R_{n2}(r)) \right]_a}{R_{n1}(a) \left[\frac{d}{dr} (r R_{n2}(r)) \right]_a - R_{n2}(a) \left[\frac{d}{dr} (r R_{n1}(r)) \right]_a}$$

$$A_2 = C(2n+1) \frac{\sigma(a)}{S} \cdot \frac{R_{n1}(a) + \mu \left[\frac{d}{dr} (r R_{n1}(r)) \right]_a}{R_{n1}(a) \left[\frac{d}{dr} (r R_{n2}(r)) \right]_a - R_{n2}(a) \left[\frac{d}{dr} (r R_{n1}(r)) \right]_a}$$

where $\mu = \beta S/\sigma(a)$.

Using Equation 16 this leads to values of M_n and N_n which differ from the values in the absence of an ocean and which will be denoted by M_n^*, N_n^* :

$$\begin{aligned} M_n^* &= \left(\frac{a}{b}\right)^2 \frac{R_{n1}(a) \left[\frac{d}{dr} (r R_{n2}(r)) \right]_b - R_{n2}(a) \left[\frac{d}{dr} (r R_{n1}(r)) \right]_b}{R_{n1}(b) \left[\frac{d}{dr} (r R_{n2}(r)) \right]_b - R_{n2}(b) \left[\frac{d}{dr} (r R_{n1}(r)) \right]_b} \\ &\quad + \mu \left(\frac{a}{b}\right)^2 \frac{\left[\frac{d}{dr} (r R_{n1}(r)) \right]_a \left[\frac{d}{dr} (r R_{n2}(r)) \right]_b - \left[\frac{d}{dr} (r R_{n2}(r)) \right]_a \left[\frac{d}{dr} (r R_{n1}(r)) \right]_b}{R_{n1}(b) \left[\frac{d}{dr} (r R_{n2}(r)) \right]_b - R_{n2}(b) \left[\frac{d}{dr} (r R_{n1}(r)) \right]_b} \\ &= M_n + \mu a^2 \frac{d}{da} \left(\frac{1}{a} M_n(a) \right) \end{aligned} \quad (2)$$

and similarly

$$N_n^* = N_n + \mu a^2 \frac{d}{da} \left(\frac{1}{a} N_n(a) \right) \quad (3)$$

These results are exact in the limit $\beta \rightarrow 0$, $\beta \rightarrow \infty$, $\beta S/\sigma(a) \rightarrow \mu$, and in fact do not depend on the assumption of constant conductivity in the oceanic shell (this is proved in the Appendix). Therefore the 'ocean' could, if required, include more highly conducting top layers of the mantle. The first-order approximation we have used is justified; even if all the material above the Mohorovicic discontinuity were included, $\beta \approx 0.005$.¹

The more general case in which the integrated conductivity of the sea varies with position is formally solved in the Appendix. This included the practical case in which the oceans do not cover the whole surface of the earth. However, for simplicity, in the discussion in §5 and §6 we assume a uniform ocean.

4. *Mathematical theory: particular solutions.* We shall now give solutions of equation 14 for our simple conductivity laws:

(a) *Constant conductivity.*

Putting

$$\sigma = \sigma_0 = \text{constant} \quad (28a)$$

we can take as basic independent solutions the functions

$$R_{n1}(r) = (r/a)^n, \quad R_{n2}(r) = (a/r)^{n+1} \quad (29a)$$

and we then find

$$R_n = (a/b)^{n+1} [n + (n+1)(b/a)^{2n+1}]/(2n+1) \quad (30a)$$

$$N_n = (a/b)^{n+2}$$

$$\cdot [1 - (b/a)^{2n+1}]/(2n+1) \quad (31a)$$

(b) *Power law conductivity.*

Putting

$$\sigma = \sigma_0(a/r)^\lambda \quad (28b)$$

where σ_0 and λ are constants, and defining

$$s_n = +[(\lambda - 1)^2 + 4n(n+1)]^{1/2}$$

we can take as basic independent solutions the functions

$$R_{n1}(r) = (a/r)^{(1+\lambda-s_n)/2}, \quad R_{n2}(r) = (a/r)^{(1+\lambda+s_n)/2} \quad (29b)$$

and we find

$$M_n = (a/b)^{(3-\lambda+s_n)/2} [(\lambda - 1 + s_n) - (\lambda - 1 - s_n)(b/a)^{s_n}]/2s_n \quad (30b)$$

$$N_n = (a/b)^{(3-\lambda+s_n)/2} [1 - (b/a)^{s_n}]/s_n \quad (31b)$$

(c) *Exponential conductivity.*

Putting

$$\sigma = \sigma_0 e^{-kr} \quad (28c)$$

where σ_0 and k are constants, we can take as basic independent solutions the functions

$$\left. \begin{aligned} R_{n1}(r) &= (2/\pi kr)^{1/2} e^{-kr/2} [\pi I_{n+1/2}(kr/2) + (-1)^n K_{n+1/2}(kr/2)] \\ &= \sum_{p=0}^{\infty} (-1)^p (n+p)!/p!(n-p)!(kr)^{p+1} \\ R_{n2}(r) &= (2/\pi kr)^{1/2} e^{-kr/2} K_{n+1/2}(kr/2) \\ &= e^{-kr} \sum_{p=0}^{\infty} (n+p)!/p!(n-p)!(kr)^{p+1} \end{aligned} \right\} \quad (29c)$$

¹ When the conductivity varies rapidly near the surface, the factor μ (which varies as $1/\sigma(a)$) would vary rapidly with the radius a at which we divided sea' and mantle. However, the terms dM_n/da and dN_n/da also vary rapidly, but as $\sigma(a)$. Thus the values of M_n^* and N_n^* are, to order β , independent of the choice of a .

where $I_\nu(z)$ and $K_\nu(z)$ are modified Bessel functions of order ν of the first and second kinds, respectively. We find

$$M_n = kb(a/b)^{3/2} e^{-k(a-b)/2} [nQ_{n,1} + (n+1)Q_{n,-1} - (2n+1)Q_{n,0}]/2(2n+1) \quad (30c)$$

$$N_n = (a/b)^{3/2} e^{-k(a-b)/2} Q_{n,0} \quad (31c)$$

where

$$Q_{n,i} = I_{n+1/2}(ka/2) K_{n+i+1/2}(kb/2) \\ - (-1)^i K_{n+1/2}(ka/2) I_{n+i+1/2}(kb/2)$$

(d) *Combined exponential and power law conductivity.*

Putting

$$\sigma = \sigma_0 e^{-kr} (a/r)^\lambda \quad (28d)$$

where σ_0 , k , and λ are constants, and defining

$$s_n = +[(\lambda - 1)^2 + 4n(n + 1)]^{1/2}$$

we can take as basic independent solutions the functions

$$R_{n1} = \sigma^{1/2} W_{k/2, s_n/2}(-kr)/r, \\ R_{n2} = \sigma^{1/2} W_{-k/2, s_n/2}(kr)/r \quad (29d)$$

where $W_{\nu, \eta}(-z)$ and $W_{-\nu, \eta}(z)$ are the confluent

hypergeometric functions known as Whittaker functions [cf., for example, *Whittaker and Watson* 1920, ch. 16]. The solutions of this subsection are included for the sake of completeness and will not be discussed further.

5. *Numerical results.* We shall now give values of M_n and N_n for several conductivity models and for $n = 1, 2, 3, 4, 5$. In each model either the conductivity has one of the simple forms a, b, c of §4 throughout the mantle, or the mantle is composed of several layers in each of which the conductivity has one of these forms. (The vectors \mathbf{Exr} and $\mathbf{j} \cdot \mathbf{r}$, and all components of \mathbf{J} must be continuous across the interfaces between such layers. These conditions are not independent and are all satisfied if R_n and Ψ_n are continuous at the interfaces.)

Table 1 defines the models themselves. (For comparison we have included a model o , in which the mantle is nonconducting.) The values of M_n and N_n are not altered if $\sigma(r)$ is increased or decreased everywhere by an arbitrary constant.

TABLE 1. Definition of Models Considered

Model	Description	Conductivity at Base of Mantle, mho/meter	Conductivity at Top of Mantle, mho/meter
o	Zero conductivity throughout mantle.	0	0
a	Constant conductivity throughout mantle.	36.0	36.0
b_1	Power law variation of conductivity, $\sigma = \sigma_0(a/r)^\lambda$ $\sigma_0 = 7.68$ mho/m, $\lambda = 5$	158	7.68
b_2	Power law variation of conductivity, $\sigma = \sigma_0(a/r)^\lambda$ $\sigma_0 = 1.14$ mho/m, $\lambda = 10$	480	1.14
c_1	Exponential variation of conductivity, $\sigma = \sigma_0 e^{-kr}$ $\sigma_0 = 1.66 \times 10^6$ mho/m, $k = 8 \times 10^{-7} \text{ m}^{-1}$	103	10.2
c_2	Exponential variation of conductivity, $\sigma = \sigma_0 e^{-kr}$ $\sigma_0 = 3.56 \times 10^8$ mho/m, $k = 2 \times 10^{-6} \text{ m}^{-1}$	338	1.04
e	Constant conductivity for the lowermost 2000 km of mantle. Zero conductivity above.	75.3	0
f_1	Constant conductivity σ_1 , for the lowermost 2000 km of mantle. Constant conductivity $\sigma_2 = \sigma_1/5$ above.	52.4	10.5
f_2	Constant conductivity σ_1 , for lowermost 2000 km of mantle. Constant conductivity $\sigma_2 = \sigma_1/100$ above.	72.4	0.145
	$\sigma_1 = 52.4$ mho/m		
g	Layer 1: $b \leq r \leq c$, $c = 1.55b = 5.39 \times 10^6 \text{ m}$ $\sigma = \sigma_1(a/r)^{\lambda_1}$, $\sigma_1 = 10.2$ mho/m, $\lambda_1 = 5.1$		
	Layer 2: $c \leq r \leq d$, $d = 1.72b = 6.00 \times 10^6 \text{ m}$ $\sigma = \sigma_2(a/r)^{\lambda_2}$, $\sigma_2 = \sigma_1(a/c)^{\lambda_1 - \lambda_2} =$ 4.0×10^{-5} mho/m, $\lambda_2 = 80$	223	0.050
	Layer 3: $d \leq r \leq a$ $\sigma = \sigma_3 = \sigma_2(a/d)^{\lambda_2} = 5.0 \times 10^{-2}$ mho/m		

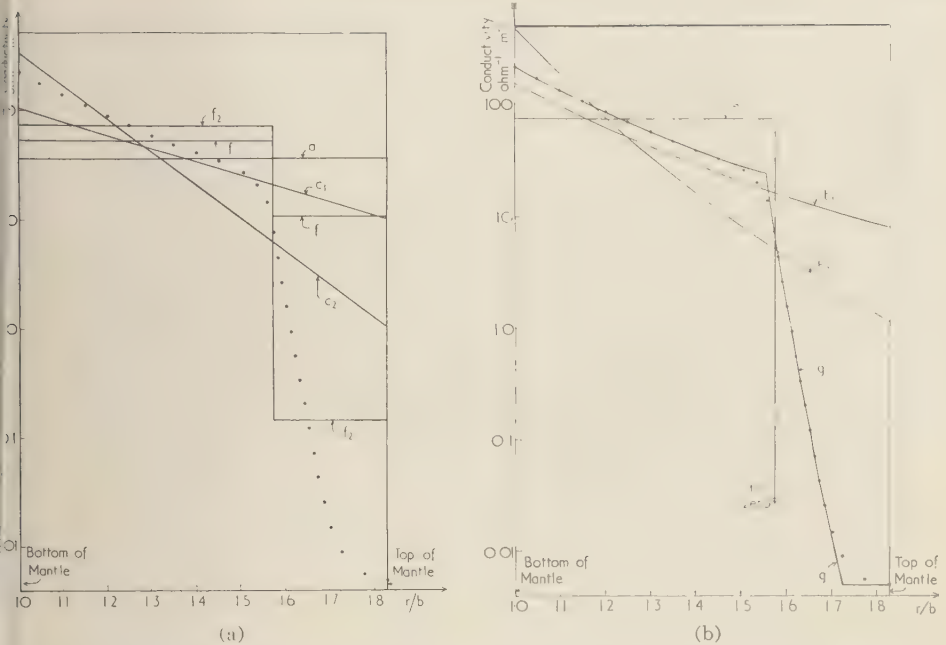


Fig. 1. Electrical conductivity distributions in the earth's mantle as defined in Table 1. The dotted curves follow that inferred by *McDonald* [1957] from a study of the screening of the secular variation by the conducting material in the mantle.

constant factor. However, for comparison, we have chosen the constant factor in each model in such a way that the electromagnetic time constant τ , of the mantle as a whole, defined by

$$\tau = \pi \left[\int_b^a [\sigma(r)]^{1/2} dr \right]^2 \quad (32)$$

is the same for each model and is equal to 3.0 years. This value is close to that implied by the distribution that *McDonald* [1957] inferred from a detailed study of the screening of the secular variation by the conducting material in the mantle.

In Figures 1a and 1b *McDonald's* distribution is shown, together with the distributions we have taken in the present analysis. In Figure 2 the variation of E_n with radius is shown for model b_1 , for $n = 1$ to 5. Figure 3 shows the variation of E_n with radius for several of the models for $n = 3$.

In Table 2, for each model the first line gives M_n and N_n ; we have taken $a/b = 1.83$. In some cases a second line gives the fractional correction to be added to M_n or N_n to allow for an ocean.

(The corrections $\delta c_1/c_1$ and $\delta c_2/c_2$ are very similar to $\delta b_1/b_1$ and $\delta b_2/b_2$, respectively, and have not been given.) The harmonic $n = 1$ corresponds to a dipole electric field, $n = 2$ to a quadrupole field, etc; there is no dependence on the azimuthal variation index m .

Model f is a two-layer model with a conducting inner shell surrounded by a poorly conducting outer shell. In f_2 the outer shell has a very low conductivity σ_2 , and practically the same M_n and N_n would be obtained in the limit as $\sigma_2 \rightarrow 0$. However these values are *not* the same as those of model e , a singular model in which $\sigma_2 = 0$. This difference can be explained as follows. When the outer layer is conducting, however slightly, a charge distribution is set up on the outside surface to prevent current flow across the surface. If the conductivity varies through the outer layer there will also be a volume charge distribution. These charge distributions do not depend very much on the value of the conductivity; they tend to a finite limit as $\sigma_2 \rightarrow 0$. However if σ_2 is put equal to zero from the start no such charge distributions are required.

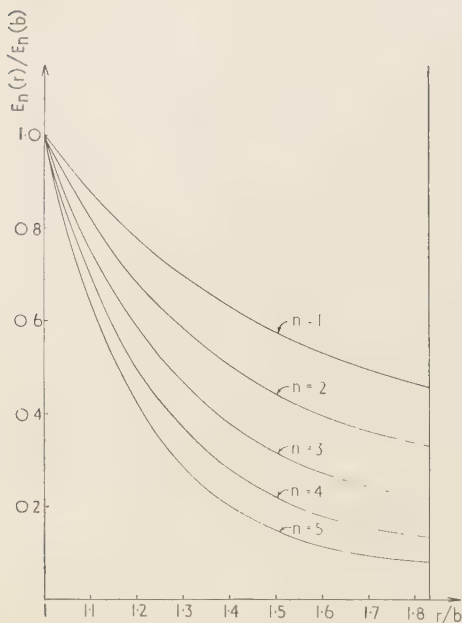


Fig. 2. Tangential electric field $E_n(r)$ in the earth's mantle for model b_1 , and $n = 1$ to 5.

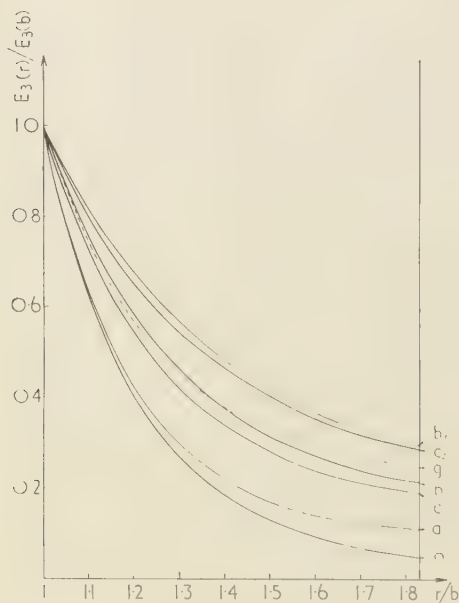


Fig. 3. Tangential electric field $E_n(r)$ in the earth's mantle for different models, and $n = 3$.

In fact model e is not applicable to the mantle as *any* conductivity of the outer shell, however small, would convert it into model f .

In discussing the values of M_n given in the table it is convenient to assume an electric field of given magnitude at the core-mantle boundary; M_n is then the ratio by which the field is reduced before it reaches the surface. In all cases the reduction is greater the higher the harmonic (see Fig. 2), but in many cases the reduction is much less than the $(a/b)^{n+2}$ which might be expected (cf. model d).

A conducting region extending from the core boundary to $r = c$ decreases M_n , the surface field increasing as c increases, to the limiting value when $c = a$, and the whole mantle is uniformly conducting (cf. distributions d , e , a). If the conductivity is then made to decrease with radius, the surface field increases still more. For the simple distributions the increase is greater the more rapidly the conductivity decreases (cf. a , b_1 , b_2 or a , c_1 , c_2); for the step-function distributions the increase is greater the greater the decrease of conductivity at the step (cf. a , f_1 , f_2).

The picture is not so simple for the more complicated distributions. The values of M_n for distribution g is less than that for b_1 (g resembles b_1 at the bottom of the mantle but has conductivity decreasing more rapidly at the top) but more than that for b_2 (which has a smaller total decrease than g but decreases more rapidly for a greater distance). However, it appears that the values of M_n and N_n for model g are typical values for any conductivity distribution which has low conductivity near the surface, a rapid rise with depth, and then a region of slow rise. This is the sort of distribution to which all present discussions of mantle conductivity lead, and so in our discussion in the next section we shall assume values for M_n ranging from 2 (for $n=1$) to 8 (for $n=5$); it is most unlikely that even an 'exotic' distribution could change these values by more than about 50 per cent. These values are very much better for detecting fields of deep origin than those 3 to 30, for uniform conductivity.

In the corrections given to allow for the effect of an ocean we have assumed that the conductivity is everywhere nonzero. If the conductivity did become zero at some radius, then any conductivity, however slight, at a larger

TABLE 2. Values of M_n and N_n

Model	M_n					N_n				
	$n = 1$	$n = 2$	$n = 3$	$n = 4$	$n = 5$	$n = 1$	$n = 2$	$n = 3$	$n = 4$	$n = 5$
o	6.13	11.22	20.52	37.56	68.73	0	0	0	0	0
a	2.71	4.81	8.97	16.8	31.3	1.71	2.13	2.89	4.16	6.24
a/a	1.26 μ	2.66 μ	3.87 μ	4.95 μ	5.98 μ	2.59 μ	3.26 μ	4.10 μ	5.04 μ	6.01 μ
b_1	2.19	3.03	4.61	7.44	12.4	0.464	0.571	0.761	1.074	1.59
b_1/b_1	0.42 μ	1.13 μ	1.98 μ	2.89 μ	3.83 μ	0.71 μ	1.31 μ	2.06 μ	2.93 μ	3.84 μ
b_2	2.04	2.51	3.34	4.77	7.15	0.220	0.260	0.328	0.439	0.617
b_2/b_2	0.22 μ	0.62 μ	1.18 μ	1.85 μ	2.59 μ	0.27 μ	0.64 μ	1.20 μ	1.84 μ	2.59 μ
c_1	2.26	3.27	5.20	8.69	15.0	0.661	0.813	1.082	1.53	2.26
c_2	2.05	2.54	3.43	4.95	7.51	0.285	0.334	0.420	0.558	0.778
e	3.09	5.18	9.28	17.0	31.5	1.52	2.01	2.81	4.10	6.83
f_1	2.42	3.79	6.46	11.4	19.2	1.24	1.52	1.98	2.76	4.02
f_2	2.34	3.54	5.83	10.02	16.2	1.13	1.36	1.76	2.41	3.47
g	2.35	2.95	4.02	5.82	8.86	0.436	0.503	0.618	0.800	1.036
$\delta g/g$	0.21 μ	0.63 μ	1.23 μ	2.00 μ	2.91 μ	0.22 μ	0.63 μ	1.24 μ	2.00 μ	2.91 μ

radius would reduce the fields to zero outside the larger radius; it is, of course, most unlikely that the conductivity ever is zero. If the conductivity did become very small not too far from the surface, we could allow for this by including with the ocean the conductivity of the higher layers.

The correction for the effect of an ocean varies nearly with $\mu = a\beta S/b\sigma(a)$, so for a given ocean it is inversely proportional to $\sigma(a)$. The percentage correction therefore depends on the absolute value of $\sigma(a)$; as stated above, we have determined $\sigma(a)$ by making all the distributions have the same time constant. In fact, with these values, and for an ocean with $a\beta S = 5000$ mms (an ocean of depth about 1 km, cf. *Lahiri and Price* [1939]), the corrections are negligible in all cases except g , which has the smallest value of $\sigma(a)$, where the maximum correction is about 5 per cent.

6. *Discussion.* The results given in Table 2 modified if necessary as our knowledge of the conductivity distribution improved) would enable us to calculate the electric and magnetic fields at the bottom of the mantle if it were possible to measure at the surface of the earth the E_θ and E_ϕ associated with these fields.

This would involve measuring steady potential gradients of continental extent against a background of many other potential distributions of widely varying magnitude, frequency, and extent. To give but two examples, there are diurnal and semidiurnal induced currents giving

fields of several mv/km (the unit in which such fields are usually measured), and many ore bodies give potential differences of about 1 volt. It would appear to be very difficult to measure fields of deep internal origin if their magnitude at the surface were very much less than 1 mv/km. This corresponds to 2 to 8 mv/km at the core-mantle boundary if we use the M_n of model g .

To give some idea of whether it would be worth looking for such fields we briefly discuss the magnitude of fields which might be expected according to present theories of the main geomagnetic field.

Most theories involve the presence of toroidal magnetic fields of magnitude of the order of 100 gauss some distance inside the core; these magnetic fields are maintained by poloidal electric currents which set up potential distributions on the core-mantle boundary. As the conductivity of the bottom of the mantle is much less than the conductivity of the core, this potential distribution is almost independent of the magnitude or distribution of conductivity in the mantle. (Putting this another way, the 'leakage' currents flowing in the mantle are very much less than the main currents flowing in the core.) The potential distribution does depend on the conductivity of the core, which we shall assume to be 3×10^6 mho/m, and on the magnitude and nature of the toroidal magnetic field. For a field of maximum value 500 gauss, with $n = 1$, and with the radial variation

given by Bullard [1949], the maximum field at the core boundary $E(b)$, is 0.065 mv/km.

For a given maximum toroidal field the value of $E(b)$ depends rather critically on the radial variation of the toroidal field. However, it proves possible to put a limit on $E(b)$ if it is assumed that there is a limit to the poloidal exciting magnetic field and to the differential rotation which together produce the toroidal magnetic field. Taking the maximum exciting field to be 4 gauss (the magnitude of the dipole field at the core boundary), and the maximum differential rotation to be 10^{-10} radian/sec (the westward drift velocity), the maximum value of $E(b)$ is 0.07 mv/km and occurs for $n = 1$.

It therefore seems most unlikely that any 'leakage' currents, associated with toroidal magnetic fields whose origin is in the core, could be detected at the surface of the earth.

In other theories, which involve thermoelectric emfs across the core-mantle boundary, the potential distribution on this boundary is the source of some of the currents in the core. If an appreciable contribution to the magnetic fields is to be made, the potential differences along the boundary have to be of the order of 1 volt [Runcorn, 1954], and this will give $E(b)$ of the order of 1 mv/km. This might lead to a measurable field at the surface of the earth.

A modification of the thermoelectric theory has been proposed by Vestine [Ingilis, 1955], who suggested that the Hall effect in the lower mantle might convert poloidal thermoelectric currents into an E-W toroidal current which gives some contribution to the dipole fields. However such Hall currents would make no direct contribution to the electric field at the surface of the earth, and the primary thermoelectric fields are unlikely to be greater than those discussed above.

We may therefore conclude that if there are any appreciable earth currents of deep origin they are most probably due to thermoelectric emfs at the core boundary, and that if such emfs make an appreciable contribution to the magnetic field inside the earth then the associated electric field at the surface of the earth may be detectable.

Acknowledgment. We wish to thank Professor S. K. Runcorn, who made the original suggestion that thermoelectric currents might be observable at the earth's surface, for his encouragement.

APPENDIX

Theory for an ocean of variable conductivity. As in the body of the paper we consider a mode of the earth in which, within the mantle ($b \leq r \leq a$), the conductivity depends only upon radial distance r from the geocenter. However, in the 'sea' ($a < r \leq a(1 + \beta)$), the conductivity is now allowed to be an arbitrary function of position (except that its variation must not be too rapid, in a sense made more precise below). The theory we develop is correct to order β .

Let D^2 be the differential operator defined by

$$D^2 = -\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) - \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} \quad (\text{A1})$$

Then (equation 9),

$$\sigma \frac{\partial}{\partial r} \left(\frac{1}{\sigma} \frac{\partial}{\partial r} (rT) \right) = \frac{D^2 T}{r} \quad (\text{A2})$$

Integrating, we find

$$\begin{aligned} rE(r, \theta, \phi) &= \frac{H_0}{4\pi\sigma(r, \theta, \phi)} \frac{\partial}{\partial r} (rT(r, \theta, \phi)) \\ &= a(1 + \beta)E(a(1 + \beta), \theta, \phi) \\ &\quad - \frac{H_0}{4\pi} \int_r^{a(1 + \beta)} \frac{D^2 T(r, \theta, \phi)}{r\sigma(r, \theta, \phi)} dr \quad (\text{A3}) \end{aligned}$$

Integrating this equation again, and using the fact that \mathbf{H} vanishes at $r = a(1 + \beta)$, we find

$$\begin{aligned} H_0 rT(r, \theta, \phi) &= -F(r, \theta, \phi) \\ &\quad \cdot \left[4\pi a(1 + \beta)E(a(1 + \beta), \theta, \phi) \right. \\ &\quad \left. - H_0 \int_r^{a(1 + \beta)} \frac{D^2 T(r, \theta, \phi)}{r\sigma(r, \theta, \phi)} dr \right] \\ &\quad - H_0 \int_r^{a(1 + \beta)} \frac{F(r, \theta, \phi)}{r\sigma(r, \theta, \phi)} \frac{D^2 T(r, \theta, \phi)}{r\sigma(r, \theta, \phi)} dr \quad (\text{A4}) \end{aligned}$$

where F is the integrated conductivity

$$F(r, \theta, \phi) = \int_r^{a(1 + \beta)} \sigma(r, \theta, \phi) dr \quad (\text{A5})$$

Thus, by (A3) and (A4),

$$\begin{aligned} H_0 T(r, \theta, \phi) &= -4\pi F(r, \theta, \phi)E(r, \theta, \phi) \\ &\quad - \frac{H_0}{r} \int_r^{a(1 + \beta)} \frac{F(r, \theta, \phi)}{r\sigma(r, \theta, \phi)} \frac{D^2 T(r, \theta, \phi)}{r\sigma(r, \theta, \phi)} dr \quad (\text{A6}) \end{aligned}$$

or small depth (i.e. $[1 - r/a(1 + \beta)]$ small) the result formally determines $T(r, \theta, \phi)$ from θ, ϕ by iteration. Provided the integral on the right-hand side of (A6) is small compared with the remaining terms, we may write

$$T(r, \theta, \phi) = -4\pi F(r, \theta, \phi)E(r, \theta, \phi) \quad (\text{A7})$$

and consequently

$$T(a, \theta, \phi) = -4\pi a\beta S(\theta, \phi)E(a, \theta, \phi) \quad (\text{A8})$$

where $S(\theta, \phi)$ is the average conductivity of the oceanic layer.

We now find conditions under which these approximations are justified. Let $\sigma_{\min}(\theta, \phi)$ be the minimum value of $\sigma(r, \theta, \phi)$ in the range $a < r \leq a(1 + \beta)$. Let $L(\theta, \phi)$ be the mean horizontal length scale of the variation of $T(r, \theta, \phi)$ in the neighborhood of (θ, ϕ) . Then

$$D^2 T(r, \theta, \phi) \approx [a/L(\theta, \phi)]^2 T(a, \theta, \phi)$$

and so the integral term on the right-hand side of (A6) is less than

$$H_0 \beta^2 S[a/L(\theta, \phi)]^2 T(a, \theta, \phi) / \sigma_{\min}(\theta, \phi)$$

which is small compared with $H_0 T(a, \theta, \phi)$ if

$$\sigma_{\min} \gg \beta^2 S(a/L)^2 \quad (\text{A9})$$

This is sufficient to ensure the validity of the approximation (A8), but, unless the variation of conductivity within the oceanic layer is extremely wild, less stringent conditions suffice. It should be noticed however that, even if the electric field at the core boundary has a mean harmonic number which is small, the mean value of $(a/L)^2$ will be large in regions where θ, ϕ varies rapidly with position. The approximation (A8) is therefore suspect in the neighborhood of oceanic margins.

Across the interface $r = a$ between the oceanic layer and the remainder of the mantle, both \mathbf{H} and $\mathbf{E} \times \mathbf{r}$ must be continuous. Hence, assuming the validity of (A8), we must solve equation 9 for the region of the mantle under the boundary condition.

$$T(a, \theta, \phi) = -\mu(\theta, \phi) \left[\frac{\partial}{\partial r} (rT(r, \theta, \phi)) \right]_a \quad (\text{A10})$$

where

$$\mu(\theta, \phi) = 4\pi\beta S(\theta, \phi)/\sigma(a) \quad (\text{A11})$$

and $\sigma(a)$ is the conductivity in the mantle immediately below the oceanic layers.

We suppose that at the core boundary the tangential electric field is $E_q^{tc} P_q^{tc}(\cos \theta) \cos \phi$. Since the system is no longer spherically symmetric because of the 'sea,' other harmonics will be excited elsewhere in the mantle. We again divide $T(r, \theta, \phi)$ into its harmonic components in a series of the form

$$T(r, \theta, \phi) = \sum_{n=1}^{\infty} \sum_{m=0}^n \sum_{j=1,2} R_{nj}(r) P_n^m(\cos \theta) \cdot (A_{nj}^{mc} \cos m\phi + A_{nj}^{ms} \sin m\phi) \quad (\text{A12})$$

At the core-mantle boundary the boundary condition on the electric field requires that

$$\frac{H_0}{4\pi b\sigma(b)} \left\{ A_{n1}^{md} \left[\frac{d}{dr} (rR_{n1}(r)) \right]_b + A_{n2}^{md} \left[\frac{d}{dr} (rR_{n2}(r)) \right]_b \right\} = E_q^{tc} \delta_{nq} \delta_{m1} \delta_{dc} \quad (\text{A13})$$

where d stands for c or s . The condition (A10) at the mantle-sea boundary requires

$$\sum_{n=1}^{\infty} \sum_{m=0}^n \sum_{j=1,2} \left\{ R_{nj}(a) + \mu(\theta, \phi) \left[\frac{d}{dr} (rR_{nj}(r)) \right]_a \right\} P_n^m(\cos \theta) \cdot (A_{nj}^{mc} \cos m\phi + A_{nj}^{ms} \sin m\phi) = 0 \quad (\text{A14})$$

The solution of (A13) and (A14) presents no theoretical difficulty. By multiplying (A14) by $P_n^{m'}(\cos \theta)_{\sin} m'\phi$ and integrating over the sphere, we can derive an infinite set of linear equations of the form

$$\sum_{n=1}^{\infty} \sum_{m=0}^n \sum_{d=c,s} \sum_{j=1,2} B_{nj}^{m'd'md} A_{nj}^{md} = 0 \quad (\text{A15})$$

where the $B_{nj}^{m'd'md}$ are constants. They are of order μ for small μ , except for the 'diagonal' elements B_{nj}^{mdmd} which are of order 1 for small μ . Thus for small μ (A13) and (A15) may be solved rapidly by iteration, and, except for A_{q1}^{tc} and A_{q2}^{tc} , each coefficient A_{nj}^{md} is of order μ ; i.e., all the extra harmonics produced by a variable sea are of order μ for small μ .

REFERENCES

- Bullard, E. C., Electromagnetic induction in a rotating sphere, *Proc. Roy. Soc. London A*, 199, 413-443, 1949.
- Chapman, S., and J. Bartels, *Geomagnetism*, Clarendon Press, Oxford, 1940.
- Inglis, D. R., Theories of the earth's magnetism, *Revs. Modern Phys.*, 27, 212-248, 1955.
- Lahiri, B. N., and A. T. Price, Electromagnetic induction in non-uniform conductors, and the determination of the conductivity of the earth from terrestrial magnetic variations, *Phil. Trans. Roy. Soc. London A*, 237, 509-540, 1939.
- McDonald, K. L., Penetration of the geomagnetic secular variation through a mantle with variable conductivity, *J. Geophys. Research*, 62, 117-141, 1957.
- Runcorn, S. K., The earth's core, *Trans. Am. Geophys. Union*, 35, 49-63, 1954.
- Whittaker, E. T., and G. N. Watson, *A Course of Modern Analysis*, 3rd ed., University Press, Cambridge, 1920.

(Manuscript received August 31, 1960;
revised January 24, 1961.)

Seismic Profiles in Northwestern Utah: Pilot Range and Grouse Creek Range Area¹

JOSEPH W. BERG, JR.,² KENNETH L. COOK, HARRY D. NARANS, JR.,³
AND RICHARD J. LEAMER⁴

*Department of Geophysics, University of Utah
Salt Lake City, Utah*

Abstract. Two seismic refraction profiles were established in northwestern Utah. The first, which extended along the area lying east of the Pilot Range, was about 55 km in total length and extended south from about 10 km southeast of Lucin, Utah, to about 13 km north of Wendover, Utah. The second, which extended north into the area west of the Grouse Creek Range, was about 95 km in total length and extended north from a point about 15 km north of Wendover to the vicinity of the town of Grouse Creek, Utah. The two profiles were reversed for about 47 km of the distance. Commercial seismic prospecting equipment, modified to pass frequencies between about 3 and 50 cps, was used.

Results of this study showed that materials having velocities of about 2.0, 5.5, 6.2, and 7.4(?) km/sec exist at depths of about 0, 0.8, 4, and 24 km, respectively. The apparent dips of the 5.5- and 6.2-km/sec horizons were about 4°N and $2\frac{1}{2}^{\circ}\text{N}$, respectively. Reflections from the horizon at an approximate depth of 24 km were identified on seismograms taken at five of the recording locations. A second reflection, which could result from a reflecting horizon at a depth of about 27 km, was identified on four of the five seismograms.

INTRODUCTION

From August 10 to September 5, 1959, seismic refraction and reflection studies were conducted in northwestern Utah. Figure 1 is an index map showing the area of investigation. The purpose of the investigation was to determine whether earth crustal studies could be conducted by means of a combination of measurements of refracted and reflected seismic pulses generated from controlled explosions of charge sizes up to a few hundred pounds.

Figure 2 is an index map of the shot points and recording stations that were occupied during this investigation. Two portable seismograph units were used which are designated in this paper as unit C and unit M. Each recording station (see explanation, Fig. 2) is designated by one of these two letters followed by the shot number. The three shot points for the investigation are shown in Figure 2 as crosses, in the

center of which is a square for shot point 1 (SP1), a single circle for shot point 2 (SP2), and concentric circles for shot point 3 (SP3). Initially, SP1, which lies about 10 km southeast of Lucin, Utah, was used as the location of the sources of seismic energy for recording stations forming a profile to the south, approximately in

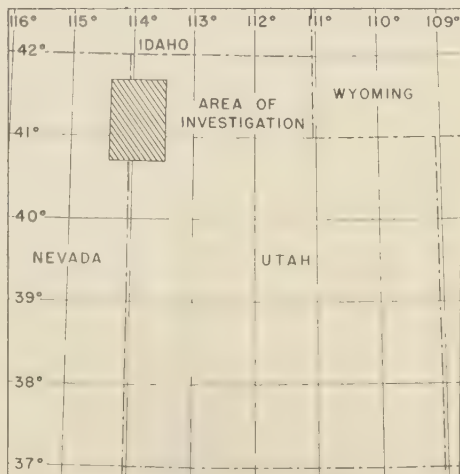


Fig. 1. Index map.

¹ Contribution No. 33, Department of Geophysics, University of Utah.

² Now at Institute for Defense Analyses, Washington, D. C.

³ Now at Shell Oil Company, Edmonton, Alberta, Canada.

⁴ Now at Mobil Oil Company, New Orleans, Louisiana.

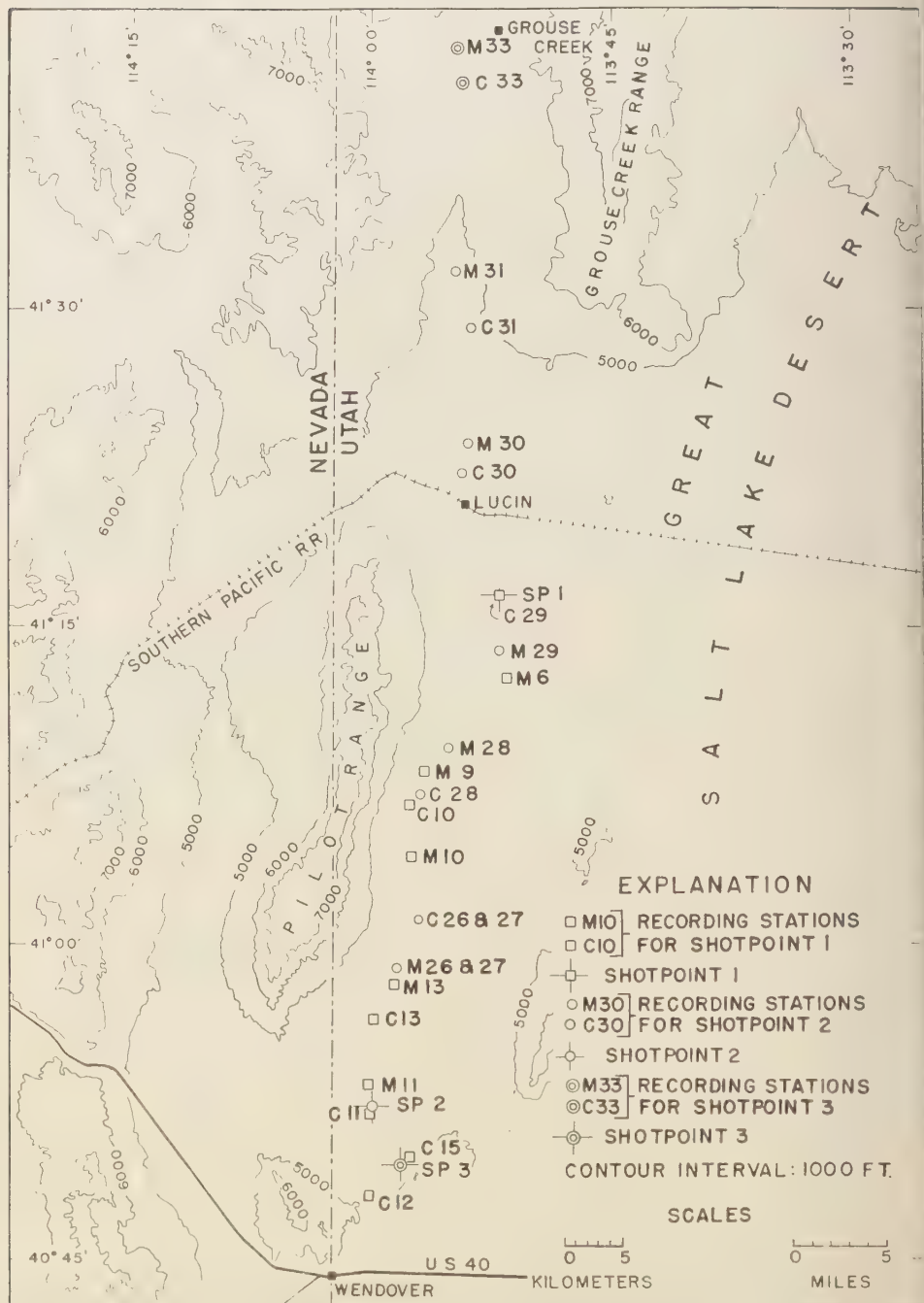


Fig. 2. Area of investigation.

line between Lucin and Wendover, Utah. The recording stations for SP1 are shown as open squares along the profile. In addition to the recording stations shown in Figure 2, a continuous refraction profile was made southward from SP1 for approximately 0.7 km, with geophone spacings of between 8 and 100 m; these stations are not shown in Figure 2. The total length of the SP1 profile (south profile) was approximately 55 km.

Stations that recorded seismic energy generated at SP2, which was about 15 km north of Wendover, are shown as open circles in a line extending north from SP2 to recording stations M31 and C31, about 20 km north of Lucin. In addition, a continuous refraction profile was made northward from SP2 for about 2.2 km, with geophone spacings ranging between 8 and 100 m; these stations are not shown in Figure 2. Stations M33 and C33, shown as concentric circles in Figure 2, recorded seismic energy from SP3, which was about 10 km northeast of Wendover, Utah. Stations that recorded energy from SP2 and SP3 constitute a north profile that is approximately 95 km in total length. The geographic coordinates of SP1, SP2, and SP3 are, respectively, $41^{\circ}16'32.9''N$, $113^{\circ}52'03.0''W$; $40^{\circ}51'20.3''N$, $113^{\circ}59'55.7''W$; and $40^{\circ}49'33.0''N$, $113^{\circ}58'09.0''W$.

PREVIOUS STUDIES

No previous seismic crustal studies have been made in the immediate area of this research, but several studies have been made in nearby areas that may have a bearing on the present results. *Patel and Tuve* [1955] conducted seismic crustal studies using explosions detonated at Kennecott Copper Company's mine at Bingham Canyon, Utah. On the basis of their refraction data, they gave a depth of about 29 km for the mantle-crust boundary and inferred a material having a velocity of 8.1 km/sec beneath this depth. *Berg, Cook, Narans, and Dolan* [1960] reported on seismic studies of nine large quarry blasts at Promontory and Lakeside, Utah, and of the Blanca underground nuclear explosion near Mercury, Nevada. The results of their refraction studies showed layers having velocities of 2.83, 5.73, 6.33, and 7.59 km/sec at approximate depths of 0, 0.2, 9, and 25 km, respectively. *Narans, Berg, and Cook* [1961] reported observing seismic reflections from layers as deep as about

25 km beneath Promontory and Rozel Point, Utah. *Schwind, Berg, and Cook* [1960] reported on seismic crustal studies involving the use of *PS* converted waves. Layering was indicated by these studies to depths of between about 25 and 31 km beneath Neola and Gold Hill, Utah, and Elko, Nevada.

GEOLOGY

The area of this investigation is in the Basin and Range province. The mountain ranges of the province are roughly parallel and trend generally in a north-south direction. Ranges 50 to 75 mi long are common and, in general, have a local relief of 3000 to 5000 ft.

According to *Butler, Loughlin, and Heikes* [1920], the Pilot Range trends a few degrees east of north and is approximately 30 mi long and 4 to 5 mi wide. Pilot Peak is the highest peak of the range, having a local relief of about 6000 ft. The range is composed of sedimentary rocks of Paleozoic age and of igneous rocks, both intrusive and extrusive. The geological structure of the Pilot Range is complex. Faulting is common and in particular there is a large east-west fault that traverses the range about midway between Pilot Peak and Lucin.

The Grouse Creek Range extends from a few miles north of Lucin, Utah, to the northern boundary of the state, a distance of nearly 40 miles. The geology of the Grouse Creek Range area is also complex, and several stocks are exposed [*Baker*, 1959]. The valleys in the area contain limestones, volcanics, and interbedded alluvial and lake sediments.

PROCEDURE

Instrumentation. The recording instrumentation used in this research consisted of vertical seismometers having a natural frequency of 2 cps (HTL S-36u), conventional seismic amplifiers modified to pass frequencies between approximately 3 and 50 cps, and conventional recording oscillographs having a paper speed of approximately 30 cm/sec. Figure 3 shows typical calibration curves for the two recording units, M and C. The ordinates shown on the calibration curves give the magnitude of ground velocity, in units of 10^{-6} in/sec, necessary to obtain a 1-inch peak-to-peak galvanometer deflection, and the abscissas give the frequency in cps. For example, on unit M in Figure 3, a ground

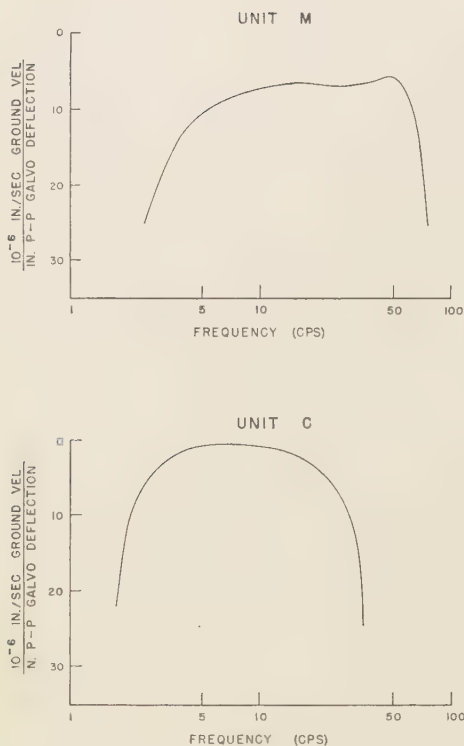


Fig. 3. Frequency response for recording units M and C.

velocity of about 10×10^{-6} in/sec with a frequency of 5 cps would result in a 1-inch deflection between the trough and crest (peak-to-peak) of the resulting trace on the seismogram.

In general, radiotelephones manufactured by the Kaar Engineering Corporation (model TR-242) were used for communication and for transmitting the shot instant to unit M recording truck. When radio transmission conditions were bad because of weather or other disturbances, time signals from radio station WWV were recorded at both shooting and recording sites to provide time control. Communications between recording units M and C were maintained with PRC-10 military-type radiotelephones. Second and minute marks from chronometers and WWV signals were recorded by both recording units (C and M) to provide time control between them. With this procedure, shot times were estimated to have an accuracy of ± 0.003 sec.

Field procedure. The shot holes were drilled in unconsolidated sediments with an auger mounted on a jeep pickup truck. The hole varied in depth from about 5 to 32 ft, averaging about 21 ft, and were $3\frac{1}{2}$ inches in diameter. A maximum of 45 lb of explosives were loaded in a hole. When larger charges were detonated, the explosive was distributed in two or more holes. At SP3, one 650-lb charge was detonated in consolidated rock about 20 ft beneath the surface of the ground in an abandoned prospect tunnel. Two types of explosive were used during the field work—60 per cent gelatin dynamite (Hercules vibrogel-B) and reclaimed 'composition B. In general, the gelatin dynamite was much better suited to shooting in the drill holes.

The electrical output from the blaster was connected to the cap, or caps, and also to the radio transmitter so as to modulate the carrier frequency at the instant of detonation. Thus the shot instant was transmitted to recording unit M. At SP3 the origin time was recorded at the shot point on a conventional recording oscillograph together with radio time signal from station WWV, and no origin time was transmitted by radio.

The seismic energy generated by the explosion was generally recorded at two receiving stations. When the shot-to-detector distance was greater than 2300 m, at least one of the two receiving stations consisted of a spread of 3 or 4 geophones placed in a line toward the shot with the distances between seismometers ranging from 150 to 600 m. With this procedure a five-man field team could detonate and record an average of about two charges a day. Charges were detonated at any time of the day, since the area of investigation was uninhabited and relatively quiet.

Reduction of data and calculations. The locations of the shot points and the recording stations were determined by triangulation from points identified on topographic maps (scale 1:250,000). Distances between the shot points and recording stations were measured directly from the maps. A check of station separations was made with the calibrated odometer of a vehicle. All distances are believed to be correct to within ± 0.1 km. Elevations were determined with altimeters tied to bench marks of known elevation and are believed to be accurate to ± 1.0 m.

Elevation corrections were made to all data to correct the travel times to the appropriate

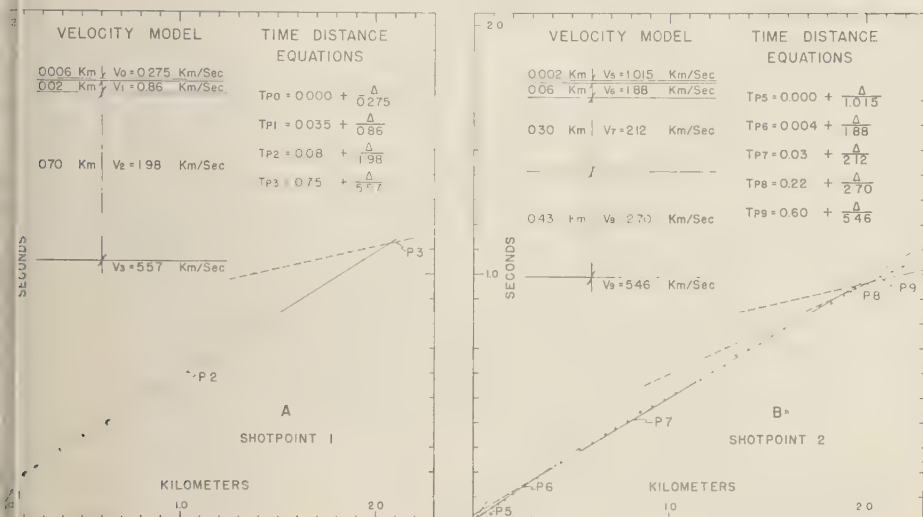


Fig. 4. Refraction profiles.

datum elevations (1326 m for profile 1 and 1303 m for profile 2); 2.0 and 5.5 km/sec were used as average velocities for unconsolidated material and consolidated rock, respectively. If a recording station was located on unconsolidated material similar to that at either shot point, the travel times were always corrected by an amount equal to the time of travel with a velocity of 2.0 km/sec along a slant path from the datum to the recording station. If a recording station was located on near bedrock, the time of travel was corrected to datum using a velocity of 5.5 km/sec. Then the time of travel with a velocity of 5.5 km/sec along a slant path from a depth of 750 m below datum elevation (which is the arbitrary depth of the 5.5-km/sec layer) to the datum elevation was subtracted from the travel time; and the time of travel with a velocity of 2.0 km/sec along a slant path across the same thickness was added to the travel time. In essence, this method was intended to simulate a uniform layering at recording locations similar to that found at the shot sites. This method of calculating elevation corrections was necessary because some of the recording stations were located on unconsolidated material similar to that found at the shot points and some were located close to bedrock. The total error in time involved in these corrections is probably within ± 0.1 sec. The elevation correction for SP3, which was located

within consolidated bedrock in an abandoned tunnel, was applied to reduce it from an elevation of 1481 m (the elevation of SP3) to the datum elevation of profile 2 (1303 m) by the last method described above.

All data reduced to datum elevations by the above methods were plotted as time-distance graphs, and straight lines were fitted to the data by the method of least squares. The depths to the refracting horizons were computed by assuming flat-lying beds and constant velocities. Next, all data from SP1 were reduced to the SP2 datum elevation. Finally, computations were made of depths and dips of the refracting horizons with the equations given by Heiland [1946, pp. 530-531]. The depths based on computations for flat-lying beds agreed to within 10 to 15 per cent with those for dipping beds.

Figure 4A shows a refraction profile to a distance of about 2300 m south from SP1, and Figure 4B shows a refraction profile to a distance of about 2300 m north from SP2. Time-distance equations are shown in the upper right-hand corners, and the corresponding velocity models are shown in the upper left-hand corners of these figures. It is believed that, in the models, the velocity of approximately 1.9 km/sec shows the depths to the water table to be 27 m and 2 m at SP1 and SP2, respectively. Drill cuttings at both shot points showed sands and silts to depths

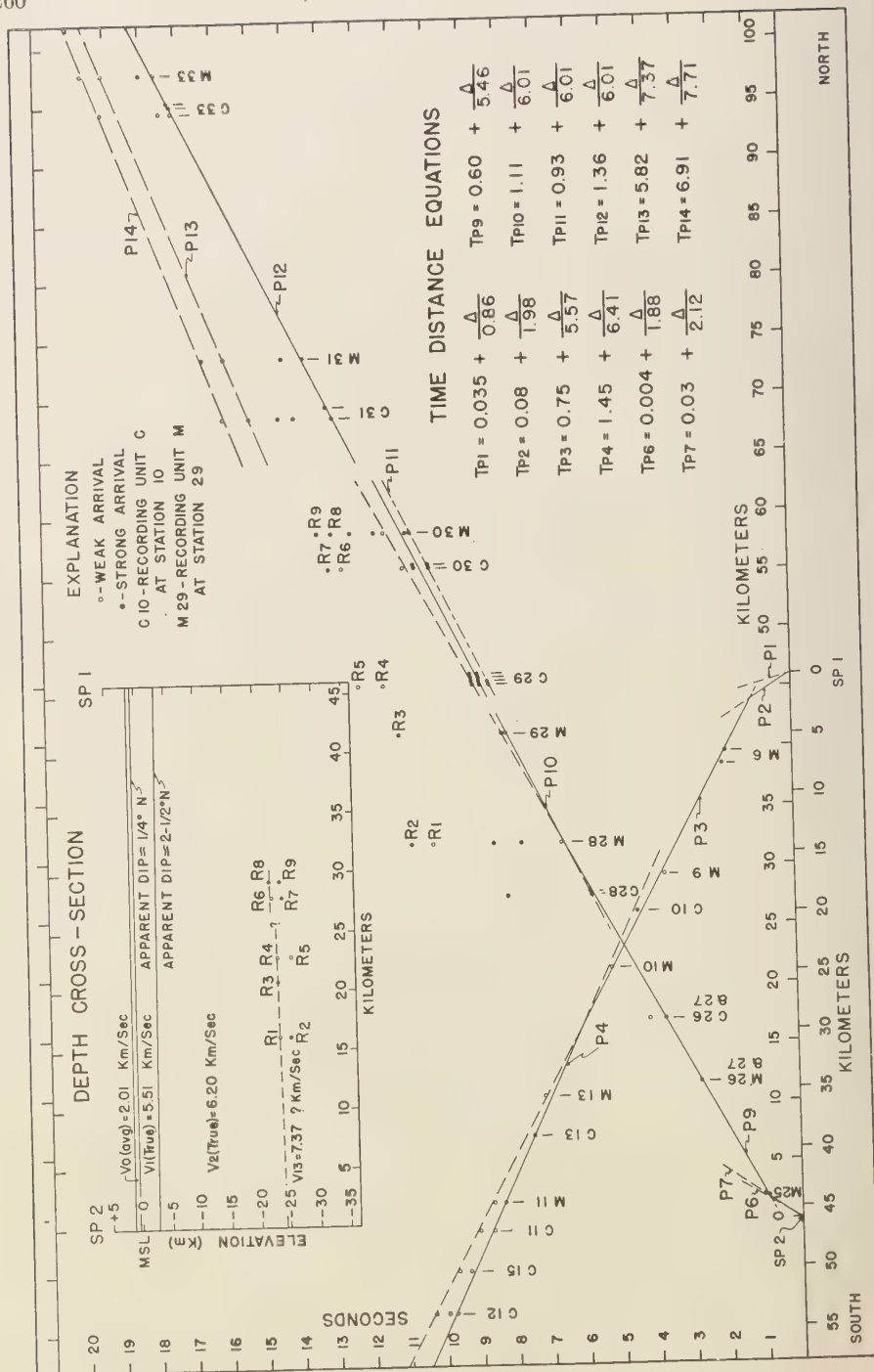


Fig. 5. Partially reversed profile for the Pilot Range and Grouse Creek Range area.

about 5 m and fine-to-coarse gravels and clay below this depth.

In Figure 4B, the horizon with a compressional wave velocity of 2.7 km/sec, at a depth of 360 m beneath SP2, is possibly composed of rocks of Tertiary age, either igneous or sedimentary. The evidence for the existence of this horizon is shown in Figure 4B as points plotted beyond a distance of 1.8 km, and P_3 represents the line fitted to these data. The data that indicate the existence of this layer could result from other geologic features, such as irregularities in the surface at the top of the 5.5-km/sec layer. There are insufficient data at SP1 to establish whether this 2.7-km/sec layer exists at this location. The results for the depth of the 5.5-km/sec layer, obtained by assuming in one instance that the 2.7-km/sec layer does exist and in the other that it does not exist, differ only 15 per cent.

PRESENTATION OF RESULTS

Figure 5 shows a partially reversed profile for the Pilot Range and Grouse Creek Range area (see Fig. 2 for locations of shot points and recording stations). Three important refracting layers are indicated by the travel-time graphs of first arrivals. Because there were small differences in velocities between comparable horizons, lines representing separate arrivals are chronologically listed, starting at SP1, as P_1 , P_2 , P_3 , and P_4 . Starting at SP2 they are chronologically listed P_6 , P_7 , P_8 , \dots , P_{14} . The time-distance equations corresponding to these lines are given in the lower right-hand corner of Figure 5. P_5 and P_{15} are omitted from the time-distance plot to preserve the clarity of presentation, but the equations are given in Figure 4B. Later arrivals, which are interpreted as reflections, are designated as R_1 , R_2 , \dots , R_8 .

Profile 1 starts at SP1 and trends south for approximately 55 km. Beneath the near-surface materials is a layer from which first arrivals are designated as P_1 on the graph and in which compressional waves travel with a velocity of about 5.6 km/sec. If constant velocities and flat-lying layers are assumed, the depth to the top of this layer is 0.7 km (see Fig. 4A). Beyond a horizontal distance of about 30 km from SP1, energy traveling with a velocity of approximately 6.0 km/sec arrives first on the seismograms, and a line fitted to these arrivals is designated as

P_4 . The depth to the surface of this horizon was computed to be about 4.7 km. The arrivals that indicated this velocity transition were weak because of poor energy coupling with the ground at SP1. All charges at SP1 were located above the water table, and first arrivals beyond a horizontal distance of approximately 30 km were weak, even though shots up to several hundred pounds were detonated. The velocity of P_4 could be in error by ± 0.15 km/sec. Later reflection arrivals could not be reliably observed on seismograms from charges detonated at SP1.

Profile 2 starts at SP2 and trends north for approximately 95 km. Designated as P_6 in Figure 5 are the 5.5-km/sec first arrivals from a layer beneath the near-surface materials. If constant velocities and flat-lying beds are assumed, the surface of this material is 0.8 km beneath SP2. First arrivals from this horizon persist to a horizontal distance of about 35 km. Beyond 35 km, the first energy to arrive travels with a velocity of 6.0 km/sec. The depth to the top of the 6.0-km/sec layer is calculated to be about 4.1 km beneath SP2.

A peculiar feature of this graph is the abrupt break in the plot in the area lying between recording positions M29 and C31. Detailed analysis was made of the data for this area, and the following interpretation is preferred. It is believed that P_{10} , P_{11} , and P_{12} are lines representing arrivals from the same horizon and that the break in the line is related to the subsurface geological structure of the rocks between recording locations M29 and C31. Figure 2 shows this area to be between the Pilot Range and the Grouse Creek Range. Apparently, a subsurface structural high, with a velocity of 6.0 km/sec, exists between the two mountain ranges. More data are needed to decipher this problem adequately.

Beyond a horizontal distance of about 68 km, later arrivals are correlated and are designated as P_{13} and P_{14} in Figure 5. If these arrivals are from a refracting surface, P_{13} has an apparent velocity of 7.4 km/sec and P_{14} has an apparent velocity of 7.7 km/sec. The data are insufficient to determine whether P_{13} and/or P_{14} actually represent refracting horizons. The critical reflection distance for the P_{13} line was computed to be approximately 72 km. For the P_{14} line, critical reflection distances of approximately 117 and 70 km were computed by assuming the existence

and nonexistence, respectively, of the 7.4-km/sec layer. The depth to the possible 7.4-km/sec horizon was computed to be 25.5 km beneath SP2. Depths to the possible 7.7-km/sec horizon were computed to be 28.9 and 25.0 km, assuming the existence and nonexistence, respectively, of the 7.4-km/sec layer. *Berg, Cook, Narans, and Dolan* [1960] showed the existence of a layer with a *P*-wave velocity of approximately 7.5 km/sec at a depth of about 25 km beneath the eastern part of the Basin and Range province. Line P_{13} could be associated with this same layer. Whether the arrivals along P_{14} are multiple or single events is not known.

Computations were made of depths and dips of the 5.5- and 6.2-km/sec horizons. The 5.5-km/sec horizon was computed to be at a depth of about 0.8 km beneath SP2 and to dip to the north at an angle of $\frac{1}{4}^\circ$. The 6.2-km/sec horizon was computed to be at a depth of about 4.1 km beneath SP2 and to dip to the north at an angle of $2\frac{1}{2}^\circ$. It is believed that the material having a velocity of 5.5 km/sec could be well-consolidated sedimentary rock, possibly of Paleozoic age, or igneous rock. The material having a velocity of 6.2 km/sec could be a crystalline basement complex.

In the upper left-hand corner of Figure 5 is an interpretive cross section which summarizes the seismic results of this investigation. Four layers are shown beneath the surface elevation of SP2

at depths of about 0.8, 4, and 24 km, and the average velocities for these layers are 2.0, 5.5, 6.2, and 7.4 km/sec, respectively. R_1 , R_3 , R_4 , R_5 , and R_8 are interpreted as reflections and are correlated to indicate the continuity of the top of the 7.4-km/sec horizon between SP1 and SP2. R_2 , R_6 , R_7 , and R_9 are plotted at depths for a weighted average velocity of 5.9 km/sec. Note that the depths computed for reflections R_2 , R_3 , R_7 , and R_9 differ only by a few kilometers from the depths of R_1 , R_5 , R_4 , R_6 , and R_8 . Thus it would appear that there is a thin layer, between 2 and 3 km thick, whose top is at a depth of about 24 km beneath the surface of the earth. The horizontal extent of this layer is not known but the reflections presented in Figure 5 indicate that it is continuous over the distance where reflections were observed. P_{14} and P_{13} could possibly be associated with the layers causing both sets of reflections. Also, the velocities of P_{13} and P_{14} could be sufficiently in error to allow P_{14} to represent arrivals from a refracting layer between 25 and 30 km below the surface.

An interesting feature of the cross section shown in Figure 5 is that the reflections indicate that the layers at a depth of about 25 km apparently dip southward. It is not known whether this is significant, but the observation warrants attention in future work.

Figure 6 shows a seismogram recorded at station C30, a distance of 55.5 km from SP2.

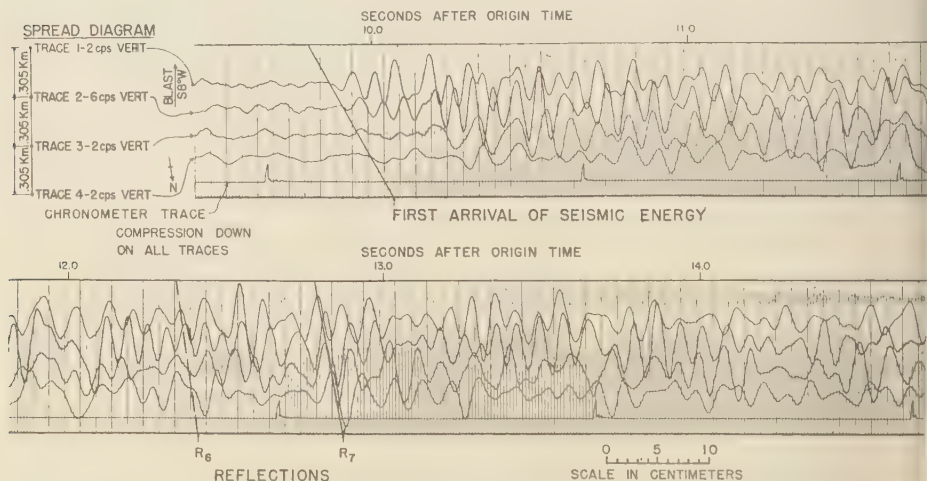


Fig. 6. Seismogram showing first arrival of seismic energy and later reflections at station C30 (charge 90 lb, 60 per cent gelatin dynamite; $\Delta = 55.5$ km).

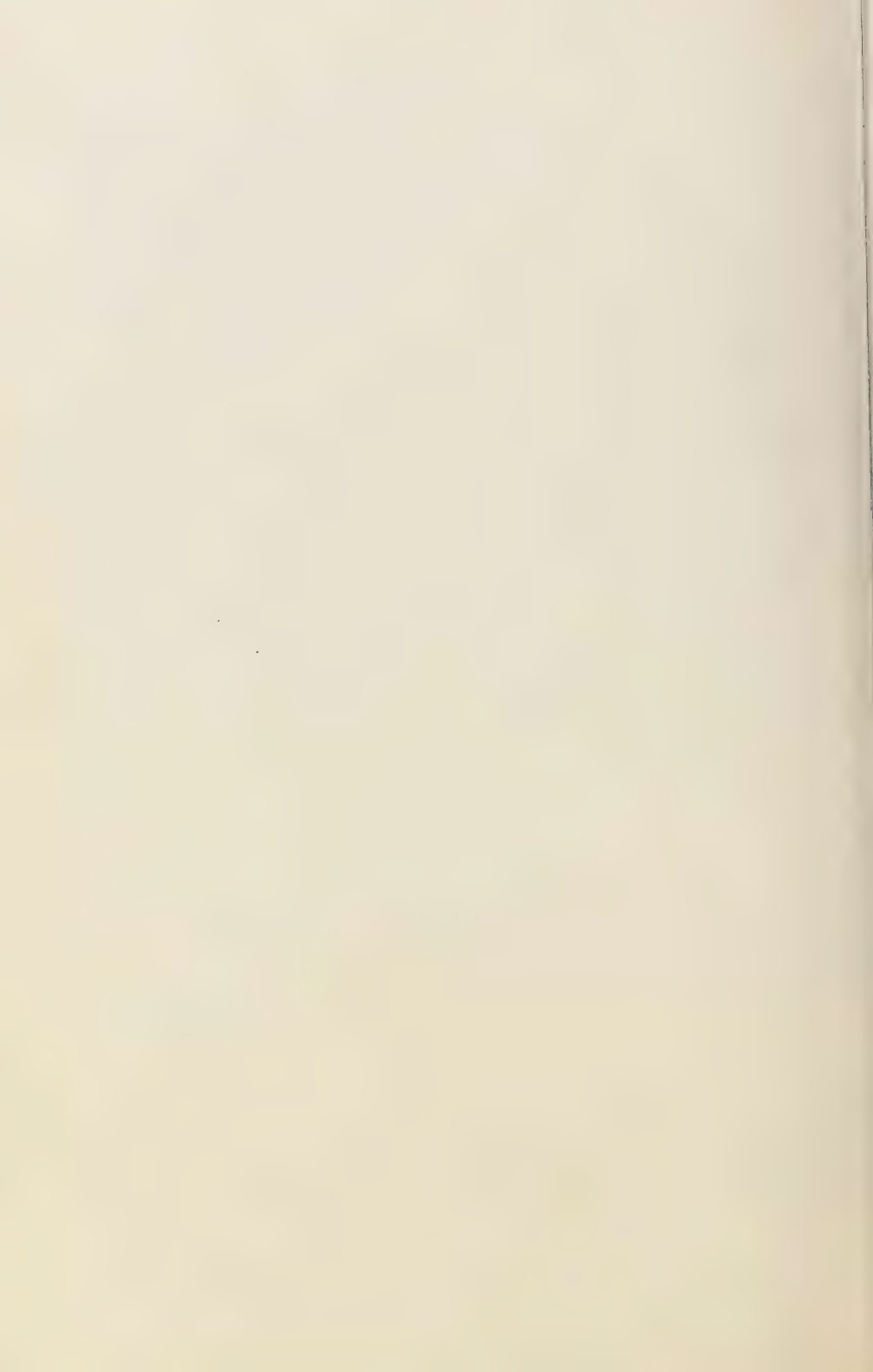
charge was 90 lb of 60 per cent gelatin dynamite. The seismometer spread for this record is shown on the left side of the figure. The first arrivals of seismic energy and reflections R_6 and R_7 which correspond to the R_6 and R_7 reflections shown in Figure 5, are indicated. Note that the reflections marked R_6 are weaker than those marked R_7 . For an average velocity of 5.9 km/sec depth to R_7 is 26 km, but we cannot be sure that the assumed velocity is correct. The reflected arrivals shown in Figure 6 are the best examples of reflections from a depth of about 25 km that were obtained during the course of this research.

Acknowledgments. Russell L. Gray and Charley Langer assisted with the field work and data reduction. During the field work, living accommodations at the Wendover Air Force Base were provided by Major Charles W. Burgess. Funds for the research were provided by the National Science Foundation under research grant NSF-G9192 and the University Research Fund of the University of Utah.

REFERENCES

- Baker, W. H., Geologic setting and origin of the Brouse Creek Pluton, Box Elder County, Utah, Ph. D. thesis, Univ. Utah, Salt Lake City, 1959.
 Berg, J. W., Jr., K. L. Cook, and W. M. Dolan, Seismic results of quarry blasts at Lakeside and Promontory Point, Utah (Abstract), *Bull. Geol. Soc. Am.*, 68 (12), 1858, 1957.
 Berg, J. W., Jr., K. L. Cook, and H. D. Narans, Jr., Seismic studies of crustal structure in the eastern Basin and Range province (Abstract), *Bull. Geol. Soc. Am.*, 70 (12), 1709, 1959.
 Berg, J. W., Jr., K. L. Cook, H. D. Narans, Jr., and W. M. Dolan, Seismic investigation of crustal structure in the eastern part of the Basin and Range province: *Bull. Seism. Soc. Am.*, 50 (4), 511-535, 1960.
 Butler, B. S., G. F. Loughlin, V. C. Heikes, and others, Ore deposits of Utah, *U. S. Geol. Survey Prof. Paper* 111, 1920.
 Heiland, C. H., *Geophysical Exploration*, Prentice-Hall, New York, 1013 pp., 1946.
 Narans, H. D., Jr., J. W. Berg, Jr., and K. L. Cook, Sub-basement seismic reflections in northern Utah, *J. Geophys. Research*, 66 (2), 599-603, 1961.
 Press, Frank, Crustal structure in the California-Nevada region, *J. Geophys. Research*, 65 (3), 1039-1051, 1960.
 Shor, G. G., Jr., Crustal structure and reflections from the Mohorovicic discontinuity in southern California, Ph.D. thesis, Calif. Inst. Technol., Pasadena, 1954.
 Schwind, J. J., J. W. Berg, Jr., and K. L. Cook, PS Converted waves from large explosions, *J. Geophys. Research*, 65 (11), 3817-3824, 1960.
 Tatel, H. E., and M. A. Tuve, Seismic exploration of a continental crust, *Geol. Soc. Am. Spec. Paper* 62, 35-50, 1955.
 Woollard, G. P., Crustal structure from gravity and seismic measurements, *J. Geophys. Research* 64 (10), 1521-1544, 1959.

(Manuscript received July 27, 1960.)



Gravity Anomalies and Crustal Section across the Tonga Trench¹

MANIK TALWANI, J. LAMAR WORZEL, AND MAURICE EWING

*Lamont Geological Observatory, Columbia University
Palisades, New York*

Abstract. In 1956, gravity observations with the Vening Meinesz pendulum apparatus were made aboard H.M.S. *Telemachus* in the southwest Pacific Ocean. For each of several gravity profiles, large free-air anomalies were associated with the Tonga and Kermadec trenches. The seismic structure determined by *Raitt, Fisher, and Mason* [1955] for the Tonga trench was projected to one of the gravity profiles. The gravity anomaly profile computed on the basis of the seismic structure is in fair agreement with the observed profile, but the two profiles differ by about 100 mgal in the vicinity of the trench. Interposing a seismically 'masked' layer under the trench and increasing the crustal thickness there to 23 km improves the fit. A crustal thickness of 36 km is derived from gravity and seismic data for the Tonga-Kermadec ridge. The disagreement between this value and one obtained by *Officer* [1955] from surface wave dispersion data is discussed.

Introduction. In the winter of 1956, a total of 32 gravity measurements were made with the Vening Meinesz pendulum apparatus aboard the submarine H. M. S. *Telemachus* in the southwest Pacific Ocean as a part of the IGY program. A report incorporating all the data obtained during this cruise will be available shortly. About sixty of the gravity stations were located in the general area of the Tonga-Kermadec trench, the Tonga-Kermadec ridge, and the Fiji ridge, structures of great tectonic and geophysical importance. In the Tonga-Kermadec trench, which lies along the edge of the southwest Pacific basin, the maximum depth sounded, $10,800 \pm 200$ meters (*Raitt, Fisher, and Mason*, 1955), is one of the greatest depths found anywhere in the world. Together with the ridge to the west, the trench and an island arc structure, typical, except in curvature, of those bordering the Pacific [*Gutenberg and Richter*, 1954, p. 47]. In particular, the region contains the epicenters of a large number of earthquakes, the deeper earthquakes occurring progressively westward, and there are many volcanic islands along the Tonga-Kermadec ridge.

The earliest marine geophysical work in this area was done by Hecker [*Heiskanen*, 1936, p. 32]. He measured the value of gravity at sea with an apparatus in which a column of mercury balanced against the pressure of a mass of air in a container at a constant temperature; photo-

graphic traces of the level of mercury in the barometer tube were obtained. The locations of his stations are given on the map (Fig. 1). For his two stations on the Tonga ridge he obtained free-air anomalies of 254 and 151 mgal; for his two stations in the Tonga trench he obtained values of -283 and -255 mgal (Table 1). Even though comparison of Hecker's values with the values obtained on H.M.S. *Telemachus* with the Vening Meinesz pendulum apparatus shows that Hecker's values are in error by about 50 to 100 mgal, Hecker was able to come to the correct conclusion that the Tonga ridge has large positive anomalies and the Tonga trench has large negative anomalies.

Members of the Danish deep sea expedition of 1950-1952, aboard R. V. *Galathea*, carried out coring and sounding operations in this area.

During the Capricorn expedition of the Scripps Institution of Oceanography (1952-1953) geophysical work which consisted of echo sounding,

TABLE 1. Hecker's Gravity Measurements on Tonga Ridge and Tonga Trench

Area	Latitude	Longitude	Depth, m	Free-Air Anomaly, mgal
Tonga ridge	28°10'S	178°27'W	2700	254
	27 15	177 40	2700	151
Tonga trench	22 07	174 13	6500	-283
	17 09	171 42	8500	-255

¹ Lamont Geological Observatory Contribution No. 483.

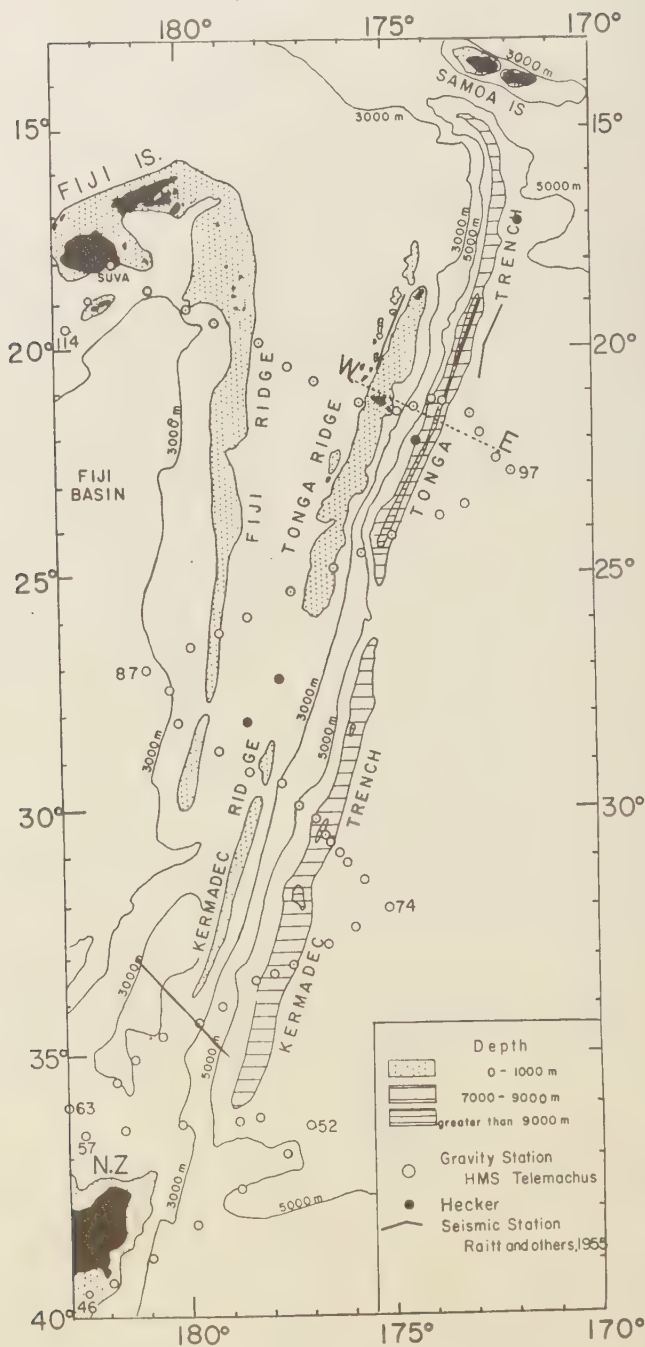


Fig. 1. Index map of the Tonga-Kermadec area.

oring, seismic refraction, and towed magnetometer studies was done in the Tonga trench area from research vessels *Horizon* and *Spencer Baird*. *Raïtt, Fisher, and Mason* [1955] constructed a bathymetric map of the Tonga trench area based on their own soundings, on the soundings obtained by R. V. *Galathea*, and on earlier charts of that area made by the U. S. Navy Hydrographic Office and the British Admiralty. From their refraction study they deduced depths to the base of the crust of 20 km below sea level in the Tonga trench, 11.5 km below sea level in the Pacific basin to the east, and at least 18 km below sea level to the west under the Tonga ridge. They also found that the basement (seismic velocity 5.2 km/sec) outcrops in the trench floor.

Officer [1955] made crustal studies in the area from an examination of the dispersion of Love and Rayleigh surface wave trains. For earthquakes which had epicenters in the vicinity of Samoa (see map in Fig. 8) and for which wave trains were recorded at Christchurch, New Zealand, the path lay almost entirely along the Tonga-Kermadec ridge. Officer compared the observed dispersion curves with theoretical dispersion curves computed for simple crustal models. For the Love waves he found that a crust 25 km thick, with mean compressional velocity between 6.0 and 6.5 km/sec, overlying a mantle with velocity of 8.10 km/sec, would fit the observations. For a fit with the observed Rayleigh wave dispersion data he found that the crust would be about 20 km thick if it had a velocity of 5.5 km/sec but would be more than 30 km thick if it had a velocity of 6.9 km/sec. (He assumed a Poisson's ratio of 0.25 in choosing shear velocities for the layers in his models.)

Brodie and Hatherton [1958] discussed the bathymetry of the Kermadec trench. They pointed out the terraced nature of the trench walls, suggesting that, at least superficially, there is some evidence of normal faulting. They also traced the extent of the shallower Hikurangi trench which lies east of New Zealand and is almost a southward continuation of the Kermadec trench.

Benioff [1949, 1954] studied the rate of release of elastic strain energy by the earthquakes in this area. One of his notable findings is that, except for a time delay of about 2000 days, there is a great resemblance between the release of elastic strain energy for the shallow earthquake

sequence of the Tonga-Kermadec region and for the intermediate earthquake sequence of the New Hebrides region to the west.

Gravity anomalies. The positions of the gravity stations occupied in the Tonga-Kermadec trench and ridge area are shown in Figure 1. A closer spacing of the gravity stations is desirable, but it was not possible because of limitations of time during the cruise. For each profile crossing the trench, attempts were made to occupy one station in the deepest part of the trench. The latitude, longitude, and depth at each gravity station, as well as the free-air anomaly and the Airy isostatic anomaly, are listed in Table 2. The positions of the stations and the free-air anomalies are also shown in Figure 2. The bathymetric contours in Figure 2 are based on the map of Tonga trench by *Raïtt, Fisher, and Mason* [1955], on the map of Kermadec trench by *Brodie and Hatherton* [1958], and on the bathymetry indicated in the General Bathymetric Chart of the Oceans, Sheets A'II and A'III (published by the International Hydrographic Bureau). Figure 3 shows the free-air anomaly versus depth for four of the profiles across the Tonga-Kermadec trench. For each of these profiles, the largest negative anomaly (about -200 mgal) coincided with the deepest part of the trench. It is interesting to compare the most negative value obtained on this cruise (-224 mgal at station 101, where the water depth was 9189 m) with the value of -346 mgal obtained at the shallower depth of 7860 m in the Puerto Rico trench [*Shurbet and Worzel*, 1956]. Gravity anomalies on the Tonga-Kermadec ridge are strongly positive (100 to 150 mgal). This is in contrast to the anomalies over the Fiji ridge, which are close to zero. The Fiji ridge is of dimensions comparable to the Tonga-Kermadec ridge (see Figs. 1 and 3), which makes the contrast between the gravity values more striking.

Two crossings, one at the south end of Kermadec trench (stations 52-57) and the other across the Hikurangi trench (stations 46-52) have not been plotted as profiles. Figure 2 shows that for both these crossings large negative anomalies are associated with the two trenches. For the crossing at the southern end of the Kermadec trench we see that the largest negative anomaly value is west of the station with the greatest water depth (station 54, -117 mgal, 6214 m; station 53, -54 mgal, 6630 m). Across the Hikurangi trench we

TABLE 2. Gravity Measurements Made Aboard H. M. S. *Telemachus* in the Tonga-Kermadec Area

Station No.	Latitude	Longitude	Depth, m	Free-Air Anomaly, mgal	Isostatic Airy Anomaly ($T = 30$), mgal
46	39 31'S	177 25'E	101	19	-14
47	39 21	178 04 E	142	62	3
48	38 51	178 59 E	3716	-42	7
49	38 14	179 54 W	3824	-4	8
50	37 34	178 49 W	3787	35	29
51	36 53	177 41 W	4480	33	28
52	36 19	177 08 W	5806	21	33
53	36 10	178 21 W	6630	-54	-24
54	36 12	178 51 W	6214	-117	-87
55	36 17	179 48 E	2914	5	3
56	36 23	178 23 E	2535	-15	-12
57	36 27	177 27 E	3588	6	52
63	35 56	177 02 E	2427	19	29
64	35 27	178 11 E	2653	60	64
65	35 00	178 40 E	2529	70	65
66	34 34	179 20 E	2396	103	80
67	34 16	179 47 W	2896	41	26
68	33 57	179 11 W	3448	4	-18
69	33 30	178 21 W	6210	-168	-153
70	33 19	177 57 W	8692	-196	-162
71	33 07	177 27 W	7106	-74	-57
72	32 43	176 36 W	6361	6	16
73	32 22	175 50 W	6148	16	24
74	31 58	175 02 W	5958	32	39
75	31 23	175 41 W	6058	26	32
76	31 02	176 04 W	5299	44	37
77	30 49	176 19 W	6318	-31	-23
78	30 38	176 30 W	7617	-137	-115
79	30 29	176 41 W	8864	-210	-179
80	30 07	176 54 W	7127	-173	-144
81	29 57	177 15 W	5854	-57	-28
82	29 26	177 37 W	3145	-7	-10
83	29 10	178 27 W	2070	72	60
84	28 42	179 11 W	2121	49	35
85	28 06	179 53 E	2690	2	-7
86	27 25	179 41 E	3414	10	17
87	26 59	179 07 E	4205	-10	8
88	26 30	179 48 W	2549	15	13
89	26 11	179 07 W	1439	40	10
90	25 53	178 27 W	2541	40	52
91	25 17	177 18 W	1531	84	64
92	24 49	176 20 W	865	127	51
93	24 29	175 40 W	4740	-108	-98
94	24 06	174 52 W	8095	-148	-123
95	23 41	173 49 W	6034	33	41
96	23 26	173 08 W	6097	23	31
97	22 44	172 02 W	6060	-3	5
98	22 29	172 21 W	6018	1	7
99	21 52	172 47 W	6460	-8	9
100	21 27	173 00 W	5938	-34	-27
101	21 09	173 38 W	9189	-224	-176
102	21 10	173 55 W	6539	-170	-134
103	21 18	174 19 W	3421	39	26
104	21 25	174 44 W	1800	153	110
105	21 10	175 41 W	720	113	55
106	20 39	176 44 W	2755	46	55
107	20 21	177 21 W	2583	28	32
108	19 51	178 02 W	2433	2	16
109	19 25	179 01 W	2147	-6	9
110	19 07	179 45 W	2872	-44	-17
111	18 40	179 22 E	2454	-51	-33
112	18 08.0	178 25.7 E	SUVA	76	
113	18 54	177 57 E	2323	-2	18
114	19 30	177 22 E	3141	20	33

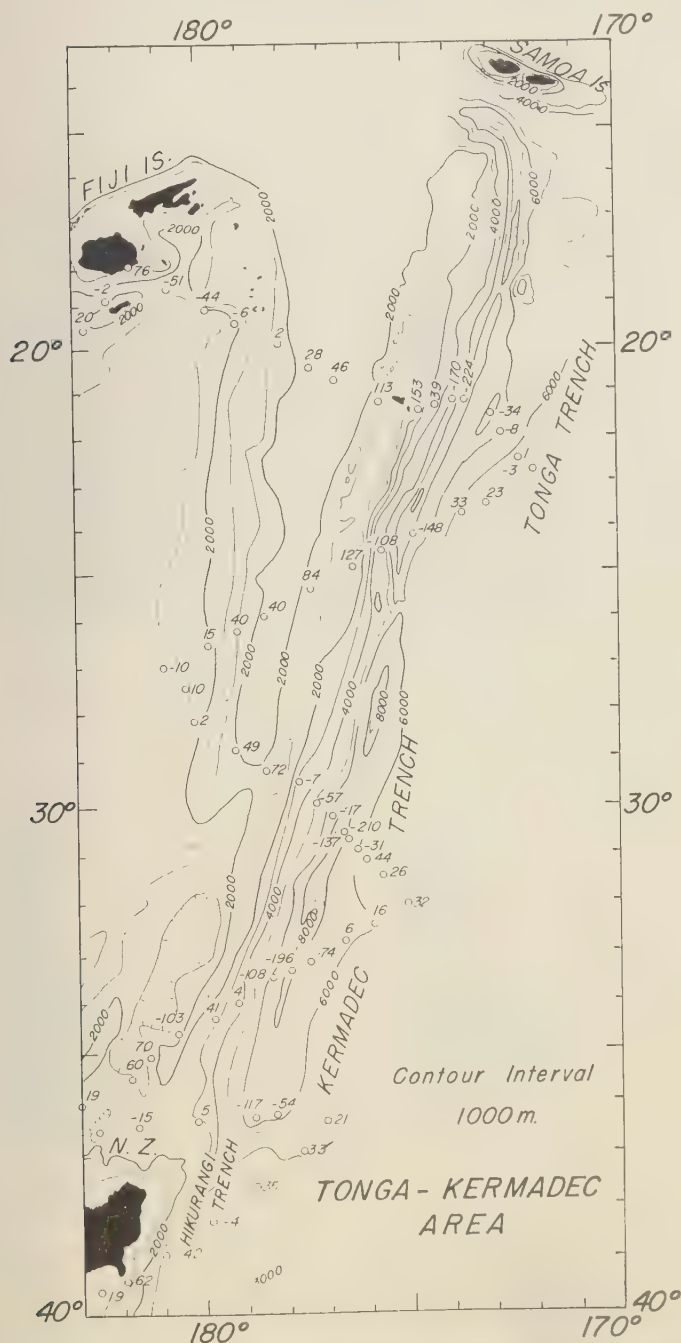


Fig. 2. Bathymetric map of the Tonga-Kermadec area. Values of free-air anomalies obtained on the H.M.S. *Telemachus* cruise are shown.

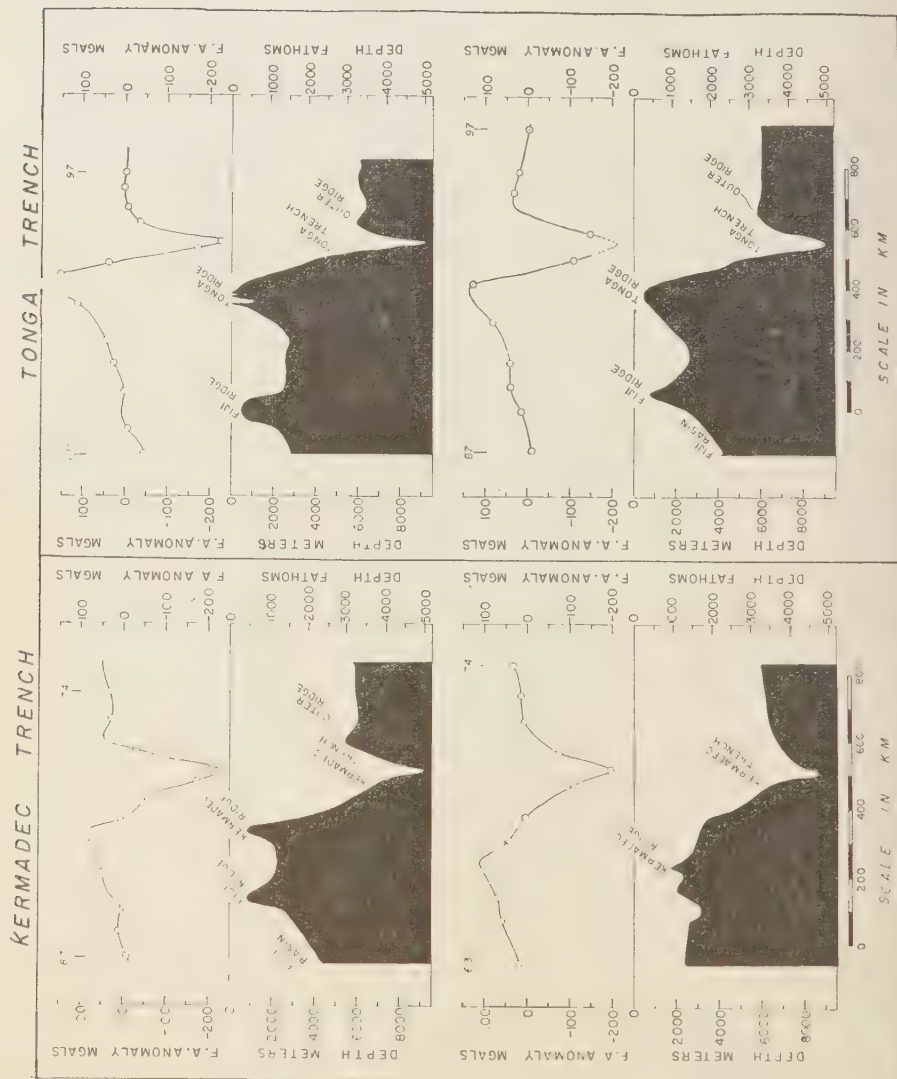


Fig. 3. Free-air anomaly and bathymetric profiles (vertical exaggeration approximately 63 to 1) across the Fiji ridge and the Tonga-Kermadec ridge and trench.

also find, for station 48, -42 mgal at 3716 m and, to the west of station 49, -4 mgal at 3824 m.

Crustal section across the Tonga trench. It is difficult to interpret gravity anomalies in great detail unless additional data, from which some independent information about the density distributions in the crust can be obtained, are available. In ocean areas such data are primarily the seismic refraction results. In the Tonga-Kermadec area only *Raitt, Fisher, and Mason* [1955] have done seismic work. They made refraction measurements along three profiles, one in the Tonga trench, one in the Tofua trough, which straddles the Tonga ridge, and one in the Pacific basin to the east. The positions of the refraction profiles are shown in Figure 1. The nearest gravity profile consists of stations 97-104. The almost perfect linearity of the trench and the ridge in this area and the fact that the bathymetric profile for *Raitt's* section is almost identical with the profile for the gravity section makes it possible to 'project' the seismic section to the line of the gravity profile (Fig. 4). To see how well the seismic section agrees with the gravity anomalies, it is necessary to assign densities to the various layers and then to compute the gravitational attraction of the whole section, which, apart from a constant to be subtracted, should be directly comparable with the free-air anomalies. The densities are assigned from an empirical curve by *Nafe and Drake* [1957] which relates seismic compressional velocities to densities (Fig. 5). The justification for using this curve is twofold. First, similar curves arrived at by different authors agree. For instance, *Wool-lard's* [1959] curve agrees closely with that of *Nafe and Drake* for densities greater than 2.5 g/cm³. Second, the crustal thickness deduced gravitationally, when this curve is used, agrees very well in areas of complex structure with that determined completely seismically. This has been demonstrated for a profile across the Puerto Rico trench [*Talwani, Sutton, and Worzel*, 1959] and another across the negative anomaly belt near Barbados in the West Indies [*Sutton, Talwani, and Worzel*, 1960]. The densities assigned from the curve are 2.0 g/cm³ for the sedimentary layer (2.0 to 2.5 km/sec), 2.6 g/cm³ for the volcanic layer (5.2 km/sec), 2.9 g/cm³ for the main crustal layer (6.5 km/sec), and 3.4 g/cm³ for subcrustal layer (8.2 km/sec).

We have assumed a uniform density (corresponding to a seismic velocity of 6.5 km/sec) for the main crustal layer, even though velocities of 7.0 km/sec and 7.6 km/sec are present under the Tonga ridge. The effect of this simplification on the gravitational results will be discussed later. To compute the gravitational attraction, two-dimensionality was assumed, and the method of *Talwani, Worzel, and Landisman* [1959] was used on the IBM 650 and the IBM 704. With this program, given any two-dimensional body represented by a polygon of arbitrary complexity, the vertical component of its gravitational attraction at any specified external point can be computed. The constant to be subtracted was obtained by specifying the computed free-air anomaly at one point. This point was chosen at 980 km (at one of the seismic refraction profiles), and the value of anomaly was taken as -21 mgal (mean of the anomalies at stations 98 and 99). The 'computed' free-air anomaly, after subtraction of this constant, is also plotted in Figure 4.

We notice that, as far as its general shape is concerned, the computed gravity anomaly curve is in fair agreement with the observed values. For stations 97, 98, and 99, at the eastern end of the profile, the agreement is very good. (This, of course, is partly due to the fact that the curve was made to fit an anomaly value of -21 mgal at 980 km). On the west side, station 104 also agrees fairly well with the computed curve. However, for the station in the trench and for those on the west flank of the trench there is considerable difference between the computed curve and the observed anomalies. For station 100, in the trench, this difference is about 85 mgal. It is extremely unlikely that this difference can be due to an error in the gravity value. An analysis of the various sources of error in the determination of gravity at sea with the Vening Meinesz pendulum apparatus led *Ewing, Worzel, and Shurbet* [1957] to conclude that the rms error in any single measurement would be about 3 mgal. This has been substantiated by measurements at secondary base stations on several Lamont Observatory cruises. We also note from Figure 3 that the value at station 100 is not unusually high when compared with the values for the other trench crossings. These barely approach -200 mgal, while the computed curve is even less than -300 mgal. Errors in navigation

are possible, but the position of station 100 would have to be in error by about 25 km for the observed anomaly to match the computed curve. Also, the computed anomaly depends on the water depth, and if the position of the station is changed the computed anomaly at the new

position changes accordingly, and the discrepancy persists. It is well known that echoes from the walls of deep sea trenches make it difficult to obtain accurate depths. Slope corrections have not been applied to the bathymetric profile. However, if this were done, the trench would

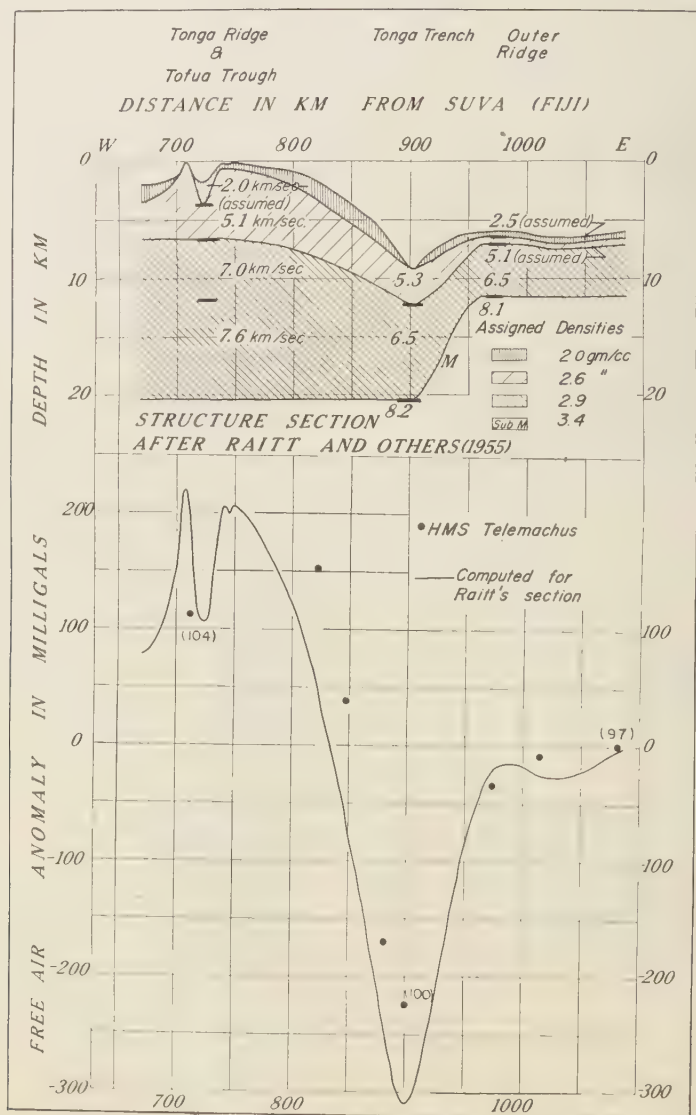


Fig. 4. Top: Structure section of Raitt, Fisher, and Mason [1955] across the Tonga trench. Densities for gravitational calculations are assigned from the Nafe-Drake empirical curve. Bottom: Computed free-air anomaly for above section. Observed values are from H.M.S. *Telemachus* cruise, stations 104 to 97.

certainly appear broader and possibly deeper, the computed anomaly would be more negative, and the discrepancy with the observed values would be larger. A different choice of densities would, of course, change the computed values, but, for reasons stated earlier, we believe that the Nafe-Drake curve is reliable. Increasing the computed anomaly in the trench by 85 mgals by raising the density of the main crustal layer requires an increase of more than 0.55 g/cm^3 . This gives a value of more than 3.15 g/cm^3 for rocks with seismic velocity of 6.5 km/sec , which seems very unusual.

Under the Tonga ridge seismic velocities of 7.0 km/sec and 7.6 km/sec have been determined by Raitt and others for the lower layers. If we assign densities from the Nafe-Drake curve to these layers (instead of assuming the same density as for the 6.5-km/sec layer) the appropriate densities would be 3.05 and 3.20 g/cm^3 . This would increase the computed anomaly by about 140 mgal , so that at station 104 the observed value would disagree with the computed value by 140 mgal . (These densities for the crust under the ridge were not used in the calculation of the curve in Figure 4 because it is difficult to calculate the attraction of a layer with a horizontal gradient.) If the crust is made appropriately thicker under the Tonga ridge, we can again make the observed and computed anomalies agree. No violation of seismic data is involved in this because, according to Raitt and others, the only restriction on the thickness of the crust here is that it be thicker than 18 km . However, under the trench the lower boundary of the crust is determined by seismic data and cannot be moved arbitrarily to make the computed curve fit the observed gravity value. For example, if a seismic travel-time curve is constructed for a thinned crust, it would not fit Raitt's travel-time observation for the trench profile.

A solution of the problem was suggested by the presence of the 7.6-km/sec layer under the ridge to the west. Figure 6 shows a travel-time curve (solid line) constructed for Raitt's model. However, an alternative model can be constructed, as is shown in the figure, by thinning the 6.5-km/sec layer and interposing a 7.6-km/sec layer between it and the mantle. Thicknesses of the 6.5-km/sec and 7.6-km/sec layers were adjusted in order to get the same intercept

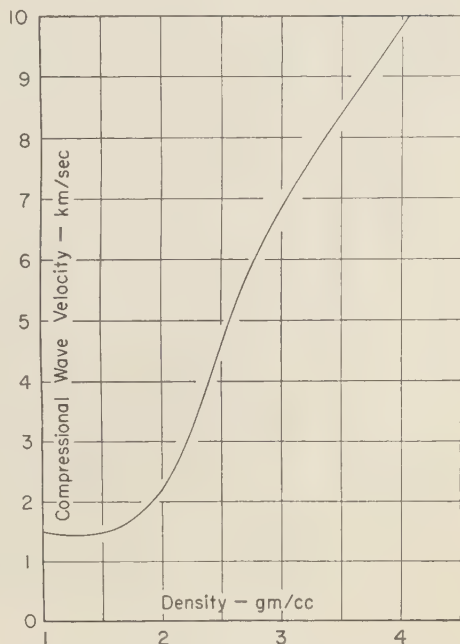


Fig. 5. Experimental relationship between compressional wave velocity and density [Nafe and Drake, unpublished data] which was used for gravity calculations.

on the time axis for the 8.2-km/sec line as in Raitt's model. Also, even though the 7.6-km/sec layer is not completely masked, its thickness is such that first arrivals from it come at about a tenth of a second earlier (dotted line in Fig. 6) than the refracted arrivals from the 6.5-km/sec and 8.2-km/sec layers. Thus it is possible that the 7.6 km/sec -layer may be present under the trench and still not be detected in an examination of the travel-time plot of first arrivals. The advantage to be gained in supposing that a 7.6-km/sec layer is present is merely that its presence makes the trench section slightly 'heavier' when a density of 3.2 g/cm^3 is assigned to the 7.6-km/sec layer from the Nafe-Drake curve, and this in turn improves the fit of the computed anomaly to the observed values.

In Figure 7 the alternative model for the crust under the trench is used. Under the Tonga ridge the 7.0-km/sec layer is again, for ease in computing, considered to be identical with the 6.5-km/sec layer in the trench. But the 7.6-km/sec layer is assigned a density of 3.2 g/cm^3 both

under the trench and under the ridge. The solid lines in the structure section in Figure 7 are obtained by joining the seismically obtained layers in the simplest fashion. Some other modifications are made to the sections of Raitt and others; for instance, the sediment is removed from the lower part of the flank of the trench. The position of the base of the crust is adjusted to get the best fit of the computed free-air anomaly (solid line) to the observed values. We notice that the fit of stations 101, 102, and 103 on the western flank of the trench is improved considerably over the fit of Raitt's section (Fig.

4). The fit of station 100 in the trench is some 35 mgal better than that of Raitt's section, but it still misses the observed value by about 5 mgal. Varying the positions of the boundaries between the layers may improve the fit of the observed values. However, in view of the sparse gravity and seismic control, such exact fitting will not necessarily determine the correct section. The section defined in Figure 7 by dotted line is one such solution. The seismic control is unchanged, and the boundaries have been shifted to improve the fit of the computed curve to the observed data. This section strikingly resembles

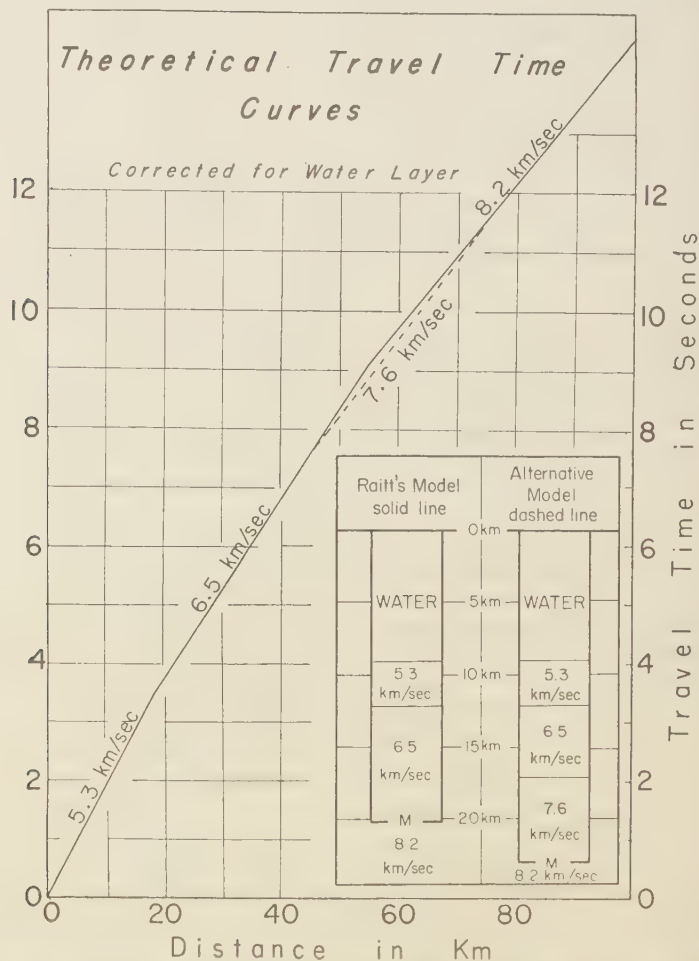


Fig. 6. Seismic refraction travel-time curves constructed for Raitt's model for the crust under the trench and for alternative model (dashed line).

the Puerto Rico crustal section [Talwani, Sutton, and Worzel, 1959].

Comparison with surface wave dispersion data. As was mentioned earlier, Officer [1955] examined

the dispersion of seismic surface waves recorded in Christchurch, New Zealand, from earthquakes in the vicinity of Samoa. The positions of the epicenters, the great-circle paths, and the Love

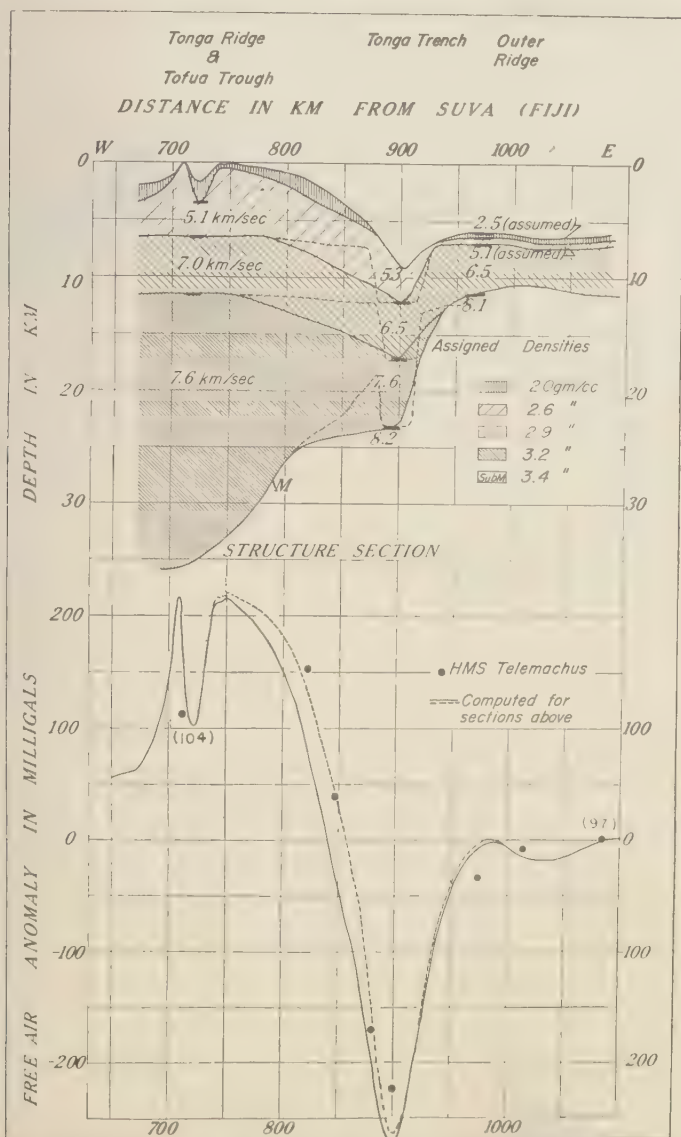


Fig. 7. Top: Structure section across the Tonga trench modified from Raitt, Fisher, and Mason, [1955] to improve the fit with the gravity data. Dotted lines show another alternative structure. Bottom: Computed free-air anomaly for above section. Dotted line corresponds to alternative structure above. Observed values are from H.M.S. *Telemachus* cruise, stations 104 to 97.

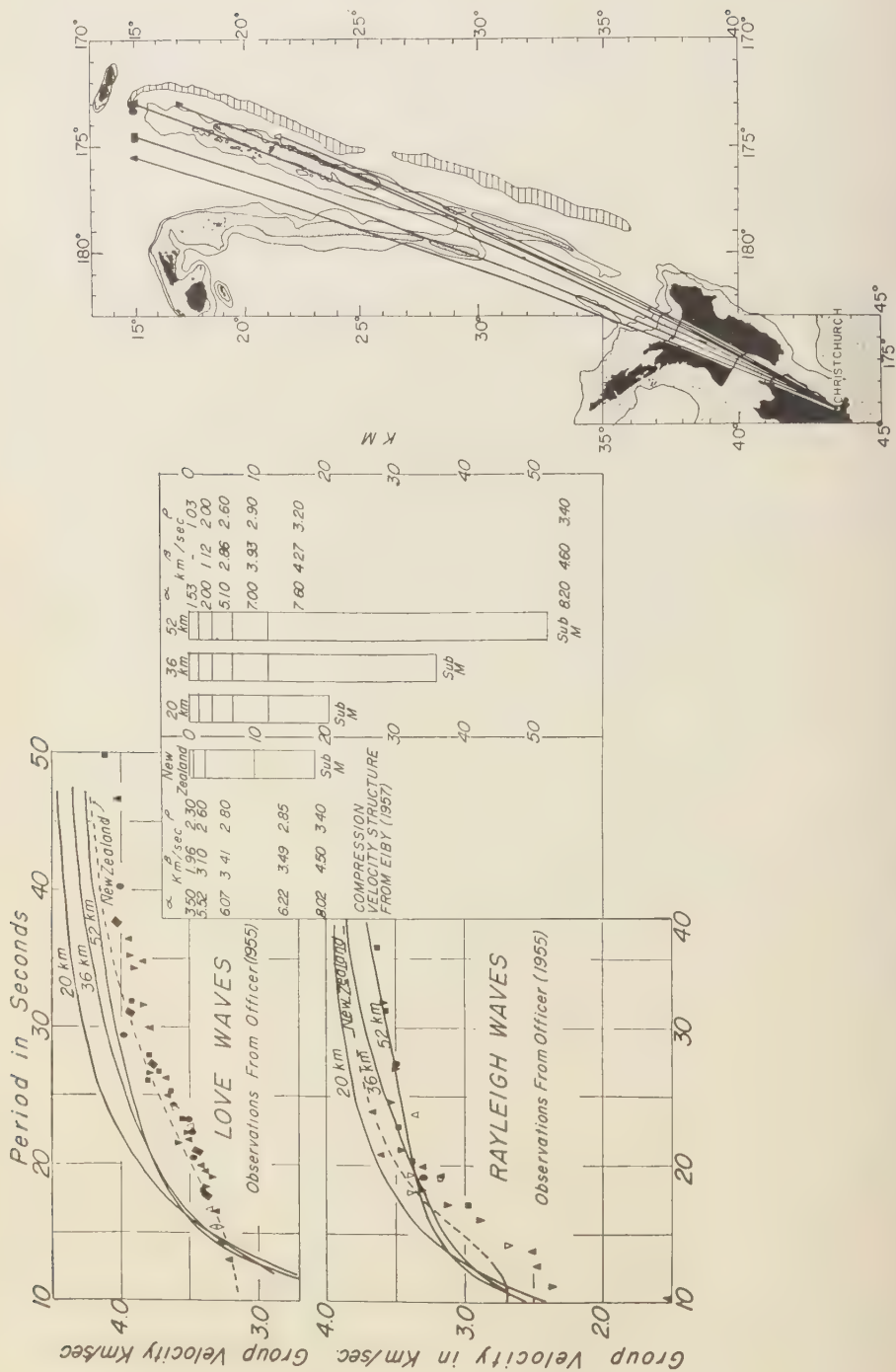


Fig. 8. Love and Rayleigh wave dispersion data along Tonga-Kermadec ridge from *Officer* [1955]. Also, computed dispersion curves for theoretical models shown.

and Rayleigh wave dispersion plots obtained by Officer are shown in Figure 8. It can be seen that a large portion of the paths for four of the earthquakes lies along the Tonga-Kermadec ridge. The programs for high-speed digital computers to obtain surface wave dispersion curves developed by Dorman [1959, 1960] enabled us to check the structure obtained from our gravity data and Raitt's seismic refraction data for the Tonga-Kermadec ridge against the observed dispersion data of Officer. For the dispersion calculations some simplifying assumptions had to be made. It was assumed that the section for the Tonga ridge arrived at from gravity and seismic considerations was typical of the entire Tonga-Kermadec ridge, with the exception of the thickness of the water layer. An average water depth of 1.5 km was chosen for the entire ridge. The compressional wave velocities were kept the same as those obtained by Raitt and others, the densities used were the same as those chosen for matching the gravity profile, and a constant Poisson's ratio of 0.27 was assumed to obtain the seismic shear velocities from the compressional velocities. This figure for Poisson's ratio was chosen on the basis of data given by Ewing and Ewing [1959, p. 301]. (They indicate an average ratio of 1.80 for compressional to shear velocity from seismic refraction measurements. This corresponds to a Poisson's ratio of 0.278. Usually, 0.25 is accepted for Poisson's ratio. As a compromise 0.27 was chosen.) Calculations for three models were made, both for the Love and Rayleigh waves. The only difference among the models was in the thickness of the lowest crustal layer which was adjusted to give total crustal thicknesses of 20 km, 36 km, and 52 km, respectively, for the three models. The models and the computed dispersion curves are shown in Figure 8. We note that for both the Rayleigh and Love waves, while the dispersion curves for the 36-km crust fit the observed data better than the curves for the 20-km crust, the fit is still far from good. Increasing the crustal thickness to 52 km improves the fit somewhat in the higher-period ranges but makes it worse in the lower-period ranges. Clearly, the solution does not lie in this direction. A part of the path for each earthquake lies in New Zealand. If the velocity of the surface waves were much lower in New Zealand than the observed velocity for the whole path, this might explain the dis-

crepancy. To check this, dispersion curves were computed for the crustal structure of New Zealand obtained by Eiby [1957]. His compressional velocities and layer thicknesses were used, and corresponding densities from the Nafe-Drake curve and shear velocities from an assumption of 0.27 for the Poisson's ratio were obtained. The computed dispersion curves (dashed lines in Fig. 8) coincide with or lie above the observed points. Thus, the mismatch between the computed curve for the ridge and the observed value for the total path (which includes the ridge and a portion of New Zealand) cannot be due to unusually low surface wave velocities in the New Zealand part of the path. The discrepancy also cannot be explained by the fact that parts of the earthquake paths lie off the ridge. Presumably, the structure off the ridge is more oceanic than that under the ridge. This would tend to increase the observed velocity, but we find that it is too low in comparison with the computed curves, not too high.

Since shear velocity is the parameter on which surface wave velocities depend most, one obvious way of matching the observed curves is by assuming lower shear velocities for the crustal layers. If we had used 3.5 or 3.6 km/sec instead of an average shear velocity of about 4.0 km/sec for the crustal layer in our models, the fit would have been better. (Officer fitted the curves by using a homogeneous crust of about 25 km thickness and about 3.5 km/sec shear velocity.) This would imply that the layers forming the crust under the ridge possess strikingly low shear velocities and a Poisson's ratio almost as high as 0.34. Detailed work is needed to prove that this is actually the case and, if so, whether this holds for other island arc ridges and mid-oceanic ridges too. Other solutions, for example those involving low-velocity layers in the crust and which do not require an abnormally high Poisson's ratio in the material underlying the Tonga-Kermadec ridge, might also reconcile the observed data with the computed curves.

Conclusions. Gravity anomalies observed over the Tonga-Kermadec trench are about -200 mgal; they are about $+150$ mgal over the Tonga-Kermadec ridge and close to zero over the Fiji ridge. The section given by Raitt, Fisher, and Mason [1955] qualitatively agrees with the gravity data. Quantitative agreement can be improved by thickening the crust under the

Tonga-Kermadec ridge to 36 km and interposing a masked layer under the trench, increasing the crustal thickness there to 23 km. Love and Rayleigh wave dispersion curves computed for the ridge structure disagree with observed values obtained by Officer [1955]. An abnormally low shear velocity in the crust underlying the ridge could explain the discrepancy.

Acknowledgments. The assistance of the officers and crew of H.M.S. *Telemachus*, Lt. Cmdr. J. E. Moore, R. N., commanding, is gratefully acknowledged. Mr. Hugh Traphagen made the observations. Mr. J. N. Rayner of the Bureau of Mineral Resources of the Commonwealth of Australia made most of the arrangements for the use of H.M.S. *Telemachus* and arranged for Mr. S. Gunson to assist in the observing program.

The calculations of Love and Rayleigh wave dispersion were made by Dr. J. Dorman of the Lamont Observatory, using his own program, on IBM 650 and 704 computers. Machine time for these as well as for the gravity computations was kindly arranged by the Watson Scientific Computing Laboratory at Columbia University.

The isostatic reductions for the H.M.S. *Telemachus* gravity stations were made by Dr. Heiskanen at the Isostatic Institute, Helsinki. Dr. L. V. Berkner helped to stimulate interest and made preliminary arrangements for the cooperative effort with the Australian and New Zealand authorities. Dr. A. Robertson of the New Zealand Department of Scientific and Industrial Research provided data and advice.

This research was jointly supported by the National Committee for the International Geophysical Year, National Academy of Sciences, under an IGY grant of the National Science Foundation, and the Office of Naval Research of the U. S. Navy.

REFERENCES

- Benioff, H., Seismic evidence for the fault origin of oceanic deeps, *Bull. Geol. Soc. Am.*, **60**, 1837-1856, 1949.
- Benioff, H., Orogenesis and deep crustal structure—Additional evidence from seismology, *Bull. Geol. Soc. Am.*, **65**, 385-400, 1954.
- Brodie, J. W., and H. Hatherton, The morphology of Kermadec and Hikurangi trenches, *Deep-Sea Research*, **5**, 18-28, 1958.
- Dorman, J., Numerical solutions for Love wave dispersion on a half-space with double surface layer, *Geophysics*, **24**, 12-29, 1959.
- Dorman, J., M. Ewing, and J. Oliver, Study of the shear-velocity distribution in the upper mantle by mantle Rayleigh waves, *Bull. Seism. Soc. Am.*, **50**, 87-115, 1960.
- Eiby, G. A., The Wellington profile, in *Crustal Structure Project*, *Geophys. Mem.* **5**, New Zealand Dept. Sci. Ind. Research, 1957.
- Ewing, M., J. L. Worzel, and G. L. Shurbet, Gravity observations at sea in the U. S. submarines *Barracuda*, *Tusk*, *Conger*, *Argonaut* and *Medregal*, in *Gedenboek, F. A. Vening Meinesz Verhandel. Ned. Geol.-Mijnbouwk. Genoot.*, *Geol. Ser.*, **18**, 49-115, 1957.
- Ewing, J., and M. Ewing, Seismic-refraction measurements in the Atlantic Ocean basins, in the Mediterranean Sea, on the Mid-Atlantic Ridge, and in the Norwegian Sea., *Bull. Geol. Soc. Am.*, **70**, 291-318, 1959.
- Gutenberg, B., and C. F. Richter, *Seismicity of the Earth and Associated Phenomena*, Princeton Univ. Press, 310 pp., 1954.
- Heiskanen, W., Das problem der isostasie, *Handbuch der Geophysik*, **1**, pt. 4, sect. 12, Berlin, 878-951, 1936.
- Nafe, J. E., and C. L. Drake, Physical properties of crustal materials as related to compressional wave velocities; paper presented at annual meeting of Soc. Expl. Geophys., Dallas, Texas, 1957. (unpublished).
- Officer, C. B., Southwest Pacific crustal structure, *Trans. Am. Geophys. Union*, **36**, 449-459, 1955.
- Raitt, R. W., R. L. Fisher, and R. G. Mason, Tonga Trench, *Geol. Soc. Am. Spec. Paper* **62**, 237-254, 1955.
- Shurbet, G. L., and J. L. Worzel, Gravity observations at sea in U.S.S. *Diabolo*, *Bull. géod.*, **42**, 51-60, 1956.
- Sutton, G. H., M. Talwani, and J. L. Worzel, West-east crustal section through the gravity anomaly belt of the Lesser Antilles along 14°20'N (Abstract), *J. Geophys. Research*, **65**, 2527, 1960.
- Talwani, M., J. L. Worzel, and M. Landisman, Rapid gravity computations for two-dimensional bodies with application to the Mendocino submarine fracture zone, *J. Geophys. Research*, **64**, 49-59, 1959.
- Talwani, M., G. H. Sutton, and J. L. Worzel, Crustal section across the Puerto Rico trench *J. Geophys. Research*, **64**, 1545-1555, 1959.
- Woollard, G. P., Crustal structure from gravity and seismic measurements, *J. Geophys. Research*, **64**, 1521-1544, 1959.

(Manuscript received November 25, 1960.)

Geomagnetic and Solar Data

J. VIRGINIA LINCOLN, EDITOR

CRPL, National Bureau of Standards
Boulder, Colorado

INTERNATIONAL DATA ON MAGNETIC DISTURBANCES

This report continues the series which has appeared regularly in this journal since vol. 54, no. 3, 295 (1949). Please refer to that first report for an explanation of the data given, and to vol. 59, no. 3, 423 (1954) for the definition of Ap.

Note: Additional and final Geomagnetic and Solar Data appear in due course in the follow-

ing international publications: *Quarterly Bulletin on Solar Activity*, International Astronomical Union, c/o Eidgenössische Steinwarte, Zurich, Switzerland; *I.G.G. Bulletins, Geomagnetic Indices K and C*, by J. Bartels, A. Romaná, and J. Veldkamp, International Union of Geodesy and Geophysics, Association of Geomagnetism and Aeronomy, c/o V. Laursen, Meteorologisk Institut, Charlottenlund, Denmark.

PRINCIPAL MAGNETIC STORMS

(Advance knowledge of the character of the records at some observatories as regards disturbances)

Observatory (Observer-in-Charge)	Greenwich date	Storm-time		Sudden commencement				C-figure, degree of activity ⁴	Maximal activity on K-scale 0 to 9			Ranges		
		GMT of begin.	GMT of ending ¹	Type ²	Amplitudes ³				Gr. day	Gr. 3-hr. period	K-index	D	H	Z
					D	H	Z							
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
College (K. Cravens)	1960	<i>h m</i>	<i>d h</i>		<i>'</i>	<i>γ</i>	<i>γ</i>					<i>'</i>	<i>γ</i>	<i>γ</i>
	Sep. 30	08 ..						ms	1	3,4,5	7	133	912	800
	Oct.		1 19											
	Oct. 2	01 ..	2 19					ms	2	3,5,6	7	270	1650	1230
	Oct. 5	22 ..	7 21					s	6	2,4,5,6	8	560	2710	1950
	Oct. 25	07 00	25 21					s	25	6	8	480	2270	1340
	Oct. 26	01 36	26 22					ms	26	3,4	7	248	1770	1250
	Oct. 28	05 30	28 23					ms	28	3,4,5	7	349	1620	1320
	Nov. 12	13 49	14 04	s.c.*	-11	-200	-65	s	13	3,4	9	1000	4260	2160
	Nov. 15	07 38	16 17	s.c.*	-95	-648	-285	s	15	5,8				
									16	1,3	8	456	3500	1500
	Nov. 21	05 ..	21 20					ms	21	3,4,6	7	265	2160	1160
	Nov. 30	19 08		s.c.*	-10	-187	-27	s	1	6	8	330	1660	1380
	Dec.		2 04											
	Dec. 6	08 ..	6 22					ms	6	4,5,7	6	200	896	838
	Dec. 15	05 ..	16 13					s	15	6	8	472	2010	2038
	Dec. 27	06 ..	27 22					ms	27	5,6	7	312	1575	885
Sitka (M. L. Clevén)	Oct. 6	03 00	9 16					s	6	3,4,5,6	9	460	2220	1340
									7	2,3,5				
	Oct. 24	14 53	31 22	s.c.*	-15	+40	+4	s	26	3	9	210	1600	950
	Nov. 4	02 00	4 17					s	4	3	8	200	880	660
	Nov. 12	13 49	16 16	s.c.*	-12	+55	+5	s	13	2,3,4,5,6	9	450	2870	2400
									16	3	9			
	Nov. 21	05 00	21 19					s	21	4	8	90	900	600
	Nov. 25	08 00	26 13					ms	25	4	7	100	600	670
	Nov. 30	19 09		s.c.*	+9	-35	-15	s	1	2,3,6	8	145	1460	900
	Dec.		2 19											
	Dec. 6	08 00	6 21					ms	6	4,5	7	45	490	430
	Dec. 15	07 00	16 13					s	15	5	8	125	1260	960
	Dec. 18	04 00	18 21					m	18	4	6	65	450	460
	Dec. 27	06 45	27 20					s	27	4	8	70	700	590

¹Approximate time of ending of storm construed as the time of cessation of reasonably marked disturbance movements in the traces; more specifically, when the K-index measure diminished to 2 or less for a reasonable period.

²s.c. = sudden commencement; s.c.* = small initial impulse followed by main impulse (the amplitude in this case is that of the main impulse only, neglecting the initial brief pulse); .. = gradual commencement.

³Signs of amplitudes of D and Z taken algebraically, D reckoned positive if toward the east and Z reckoned positive if vertically downwards.

⁴Storm described by three degrees of activity: m for moderate (when K index as great as 5); ms for moderately severe (when K = 6 or 7); s for severe (when K = 8 or 9).

PRINCIPAL MAGNETIC STORMS. Continued

Observatory (Observer-in-Charge)	Green- wich date	Storm-time		Sudden commencement				C- figure, degree of ac- tivity ⁴	Maximal activity on K-scale 0 to 9			Ranges		
		GMT of begin.	GMT of ending ¹	Type ²	Amplitudes ³				Gr. day	Gr. 3-hr. period	K- index	D	H	Z
					D	H	Z							
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Witteveen (D. v. Sabben)	1960 Oct. 4	<i>h m</i> 14 00	<i>d h</i> 9 24	s	6	7,8				
									7	1	9	135	1215	839
	Oct. 24	12 52	31 24	s.c.*	-5	+47	+2	s	25	6	9	65	530	240
	Nov. 3	20 00	4 24	ms	6	6	6	30	140	125
	Nov. 12	13 49	14 21	s.c.*	-4	+57	-2	s	13	1,3,4,5	9	190	1085	725
	Nov. 15	13 04	16 17	s.c.*	-4	+52	+2	s	16	1,2	8	60	600	480
	Nov. 21	06 32	22 07	s.c.	+1	+12	0	ms	21	7				
									22	1	6	40	195	125
	Nov. 30	19 10		s.c.	-2	+104	0	ms	1	2	7	65	260	200
	Dec.		2 02											
	Dec. 7	18 04	8 11	s.c.*	-2	+38	0	ms	7	8	7	45	99	57
	Dec. 15	13 00	16 14	ms	15	5,6,8	6	60	205	210
	Dec. 27	03 00	28 20	ms	27	6,7	7	45	275	115
	Fredericksburg (R.E.Gebhardt)	Sep. 30	13				m	1	2,3	5	29	135
			2 19						2	3,4,5				
Oct. 4		14 ..	5 10				m	4	7,8	5	26	110	64
Oct. 5		19										
Oct. 6		02 39	10 05	s.c.	+2	+19	-5	s	7	1	9	108	640	651
Oct. 24		14 52		s.c.*	-2	-42	-3	ms	25	6	6	41	237	159
			1 09						26	1				
									28	4				
Nov. 3		22 27	5 09	s.c.*	+3	+20	-1	ms	4	1	6	31	113	82
Nov. 12		13 49		s.c.*	-6	-69	-4	s	13	3,4,5	9	219	1343	858
Nov. 15		13 04	17 21	s.c.*	-3	-40	+6	ms	15	8	7	49	285	226
									16	1	7			
Nov. 21		03 ..	22 13				m	21	4	5	26	124	64
									22	1,2				
Nov. 24		12 33		s.c.*	0	+2	0							
Nov. 24		20 53	26 15	s.c.*	+1	-27	+3	m	24	8	5	31	99	6
									25	2,3	5			
Nov. 30		19 09		s.c.*	+3	+40	-7							
Nov. 30		23 58		s.c.*	+5	+30	-6	ms	1	2	7	52	182	126
Dec.			2 24											
Dec. 7		18 04	8 16	s.c.*	+3	-17	-4	m	8	1,2	5	22	65	54
Dec. 14		24 ..	16 16	m	15	8	5	26	168	116
								16	1,4					
Dec. 27	03 ..	28 21	m	27	3	5	26	141	77	
Tucson (C. J. Beers)	Sep. 29	21										
	Oct.		3 15					ms	2	4	6	17	118	31
	Oct. 4	16 ..	5 11				m	4	6,7,8	5			
									5	3		14	110	36
	Oct. 5	19 ..	10 04				ms	6	2,7				
									7	1,2,3	7	38	426	137
	Oct. 24	14 52	31 23	s.c.*	+3	+21	+3	ms	25	6	7	23	204	51
	Nov. 12	13 50	14 12	s.c.*	+5	+33	+2	s	13	3,4	9	87	465	117
	Nov. 15	13 04	17 11	s.c.	-3	+6	ms	15	8				
									16	1	7	24	201	60
	Nov. 21	04 ..	22 18				ms	21	4	6	18	102	26
	Nov. 24	20 52	26 15	s.c.*	+1	-21	-1	ms	25	4	6	18	99	26
	Nov. 30	19 09		s.c.*	+1	+9	+1	ms	1	3	7	28	210	4
	Dec.		2 23											
	Dec. 7	18 04	8 15	s.c.*	+2	+9	m	7	7,8	5	15	69	3
									8	1				
	Dec. 15	07 ..	16 14	m	15	5,6	5	18	164	49
								16	1,4					
Dec. 27	03 ..	28 23	m	27	3,4,6					
								28	4	5	17	124	37	
San Juan (M. Vazquez)	Oct. 5	19 ..	8 03				ms	6	7	7	20	333	60
									7	1,2				
	Oct. 24	14 52	26 15	s.c.	-1	+12	-5	ms	25	6	7	8	231	37
	Nov. 12	13 25	14 12	s.c.	+1	+20	-3	s	13	3	8	46	366	15
Dec.	None												
Honolulu (E. K. McCaffrey)	Sep. 30	18 00					m	1	3	5	5	85	1
	Oct.		2 21						2	3				
	Oct. 4	15 00	8 12				ms	6	8	7	10	300	49
									7	3				

PRINCIPAL MAGNETIC STORMS. Continued

Observatory (Observer-in-Charge)	Greenwich date	Storm-time		Sudden commencement				C-figure, degree of activity ⁴	Maximal activity on K-scale 0 to 9			Ranges		
		GMT of begin.	GMT of ending ¹	Type ²	Amplitudes ³				Gr. day	Gr. 3-hr. period	K-index	D	H	Z
					D	H	Z							
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Honolulu (Continued)	1960	<i>h m</i>	<i>d h</i>			<i>γ</i>	<i>γ</i>						<i>γ</i>	<i>γ</i>
	Oct. 13	21 47	14 03	s.c.*	0	+22*	+10*	m	13	8	4	0	30	10
	Oct. 24	14 52	30 14	s.c.	+1	+24	+9	ms	25	3	6	5	135	20
	Nov. 12	13 49	17 12	s.c.	-1	+30	+15	s	13	3,4	8	43	410	130
	Nov. 20	23 00	22 18	m	21	6	5	2	80	10
	Nov. 25	17 07	26 15	m	24	8	5	4	120	20
									25	4				
	Nov. 30	19 10		s.c.*	+2	+22	+8	ms	1	1,2	6	5	215	48
	Dec. 7		2 24											
	Dec. 7	18 05	8 16	s.c.*	+1	+20	+2*	m	8	1	5	2	42	10
	Dec. 14	09 00	16 12	ms	16	4	6	3	34	6
	Dec. 25	20 03	28 16	s.c.	+1	+9	+2	m	27	3,7	5	3	40	15
									28	4	5			
	Huancayo (A. Giesecke) (N. Casaverde)	Oct. 4	15 35	5 06	ms	4	6,7	7	9	433
Oct. 6		02 40	7 23	ms	6	5,6,7,8	7	13	607	66
Oct. 24		14 52	26 17	s.c.*	+4	+185	+12	ms	25	6,7	7	12	633	43
Nov. 3		16 20	4 20	m	4	5,6	5	8	260	40
Nov. 11		00 33	11 22	s.c.	0	+14	+3	m	11	4,5,6,7	5	8	223	56
Nov. 12		13 22	14 21	s.c.	+63	s	13	5,6	9	(30)	924	117
Nov. 15		13 06	16 21	s.c.	+2	+66	+2	ms	15	6,8	7	11	448	47
Nov. 21		05 32	22 03	ms	21	7	7	9	401	25
Nov. 30		19 09		s.c.*	+4	+133	+9	ms	1	5,6,7	6	13	317	42
Dec. 7			1 24											
Dec. 7		18 04	8 17	s.c.*	+3	+123	+7	m	8	5,6	5	7	287	40
Dec. 15		11 03	16 14	ms	15	6	6	7	214	45
Dec. 18		03 13	19 03	ms	18	6	6	10	218	49
Dec. 27		07 12	27 21	ms	27	6	7	10	283	55
Isabellville (A. M. van Wijk)		Oct. 6	02 40	7 21	s.c.*	+20	s	6	4,8		20	(1)500
	Oct. 24	14 52	25 21	s.c.	+0	+26	-2	m	25	6,7		10	310	28
	Nov. 12	13 47	13 24	s.c.	+8	+37	-4	s	12	7,8		46	(1)500	49
	Dec. 7								13	1,7				
Dec. No important storms														
(1) The "Ranges" H of October 6 and November 12 are approximate in consequence of defective motion of the repeating prisms.														
Lima (G. Keys)	Oct. 1	08 34	3 02	m	1	3	5	2	95	20
	Oct. 4	14 ..	10 03	ms	6	8	7	10	368	67
	Oct. 24	14 52		s.c.	0	+25	-8	ms	25	3	6	8	151	39
	Nov. 12	13 48	14 13	s.c.	0	+35	-16	s	13	4	8	23	466	151
	Nov. 15	02 20	17 12	s.c.	+1	+29	-15	ms	16	1	6	13	230	53
	Nov. 21	06 31	23 07	s.c.	0	+13	-6	m	21	6	5	8	105	32
	Nov. 30	19 08		s.c.	ms	1	2	7	14	240	43
	Dec. 7		2 21											
	Dec. 7	18 03	10 02	m	8	1	5	10	112	24
	Dec. 13	22 ..	16 17	ms	16	4	6	8	200	35
	Sep. 29	08 37		s.c.?	0	+3	0	m	30	7	5	22	123	77
	Oct. 4	14 ..	5 02	ms	4	8	6	22	115	137
	Oct. 5	19 ..	8 12	s	6	8	8	59	318	329
	Oct. 9	21 ..	10 00	Bay	m	9	8	5	12	43	55
	Oct. 15	03 ..	16 00	m	15	6,7	5	21	134	99
Oct. 24	14 52	24 21	s.c.*	+3	+20	+16	ms	24	6	7	17	157	140	
Lima (A. M. van Wijk)	Oct. 25	02 ..	31 00	ms	25	7	7	35	222	204
	Nov. 11	00 34	12 00	s.c.?	-1	+10	+5	m	11	4	5	25	113	83
	Nov. 12	13 48	14 21	s.c.	+4	+30	+23	s	13	2,7	8	70	451	304
				Strong S.F.E. at 1324, Nov. 12.										
	Nov. 15	13 04	17 00	s.c.	+1	+14	+9	ms	15	8	7	38	204	164
	Nov. 21	06 32	22 09	s.c.?	+2	+13	+4	ms	21	7	6	25	171	119
	Nov. 30	19 10		s.c.	+1	+30	+17	ms	1	3,7	6	28	200	134
	Dec. 7		2 22											
	Dec. 6	18 ..	6 22	Bay	m	6	7	5	10	69	89
	Dec. 7	18 04	8 16	s.c.	+1	+21	+7	m	7	7,8	5	16	88	85
									8	2	5			
	Dec. 15	13 ..	16 14	ms	15	7	6	28	110	129
	Dec. 17	23 ..	18 21	m	18	3,4,5,6	5	17	121	187
	Dec. 25	20 02	28 20	s.c.	+1	+13	+6	ms	27	7	6	23	180	121

PRINCIPAL MAGNETIC STORMS. Continued

Observatory (Observer-in-Charge)	Greenwich date	Storm-time		Sudden commencement			C-figure, degree of activity ⁴	Maximal activity on K-scale 0 to 9			Ranges			
		GMT of begin.	GMT of ending ¹	Type ²	Amplitudes ³			Gr. day	Gr. 3-hr. period	K-index	D	H	Z	
					D	H								Z
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Gnangara (P. M. McGregor)	1960	<i>h m</i>	<i>d h</i>		<i>'</i>	<i>γ</i>	<i>γ</i>					<i>'</i>	<i>γ</i>	<i>γ</i>
	Sep. 30	00				ms	1	5	6	23	92	16
	Oct.		2 20					2	4				
	Oct. 4	14 45	9 16				s	6	8	8	47	327	33
	Oct. 24	14 52	13 21	s.c.	+3	+36	+11	ms	25	5,6	7	34	183	20
	Nov. 3	20 ..	5 00				m	4	1,3,5,6	5	19	108	11
	Nov. 11	00 32	12 01	s.c.*	+2*	+3*	+6*	ms	11	4	6	20	107	13
	Nov. 12	13 48	15 00	s.c.	+4	+56	+22	s	13	1,3,4	9	94	390	660
	Nov. 15	13 05	17 20	s.c.	+1	+32	+8	ms	15	8	7	47	375	16
									16	1,2				
	Nov. 21	06 32	22 19	s.c.	+5	+22	+20	ms	21	7	6	29	100	14
	Nov. 30	19 08		s.c.*	+6*	+31	+32*	ms	1	1,3,5	6	37	155	19
	Nov. 30	23 58		s.c.*	-7*	-23*	-27*							
	Dec.		2 20										
Toolangi (C. A. van der Waal)	Dec. 15	02 ..	16 17				ms	15	5,8	6	22	147	160
	Dec. 27	00 ..	28 20				ms	27	6	6	27	109	13
	Oct. 6	02 39	7 18	s.c.	0	+27	+2	s	6	8	7	77	>370	16
				Reserve trace on H component illegible from 23:00Z till 24:00Z.										
	Oct. 24	14 52	26 18	s.c.	+1	+54	+10	ms	24	6	6	29	229	160
									25	3,5,6,7				
									26	4				
	Nov. 12	13 48	s.c.*	-3*	+82	+13	Progress of storm obscured by new commencement at 18:46Z						
	Nov. 12	18 46	14 11	s.c.	+12	+61	+6	s	12	8	8			
									13	2,3,4	8	63	660	42
	Nov. 15	22 ..	16 16				s	15	8	8	30	398	11
	Nov. 30	19 09		s.c.*	+3*	+21	+3	ms				35	232	11
	Dec.		2 06					1	2,3,4	6			
	Dec. 7	18 04	8 11	s.c.*	+2*	+19	+4	m	7	7	5	22	130	3
	Dec. 15	07 30	16 13				m	15	5	6	33	170	12
	Dec. 27	00 00	28 16				m	27	6,7	6	28	166	7
Mawson Antarctica (M. Kirtton)	1950													
	Feb. 11	07 58	11 16	s.c.*	-11*	+110*	ms	11	4	7	132	930	5
	Feb. 16	16 19	16 23	s.c.	-90	-135	ms	16	7	7	153	1316	...
	Feb. 25	02 50	26 10	s.c.	-21	+85	+70	ms	25	2,3,4,6	7	184	1215	(900)
	Mar. 26	08 40	27 18	s.c.*	-7	+122*	+25	ms	26	5,6,8	7	130	1008	94
									27	1				
	Mar. 28	18 17	29 22	s.c.	-7	-292	-35	ms	28	7	8	167	(1900)	153
									29	1				
	Apr. 10	03 11	10 22	s.c.	+6	-108	-65	m	10	4	7	102	540	82
	May 11	00 53	12 24	s.c.	+4	-45	-15	ms	12	6	8	232	2120	154
	Jul. 11	16 23	12 04	s.c.*	+9*	-278*	-90*	m	11	6,7,8	7	1015	1375	119
									12	1				
	Jul. 15	08 06	16 02	s.c.	-13	(+150)	(+120)	s	15	4,5,8	8	250	2460	179
	Jul. 17	16 42	19 11	s.c.	-13	+133	-95	s	17	7	8	207	1560	159
	Aug. 16	04 00	17 12				ms	16	6,8	7	(215)	(1970)	141
									17	1,2				
	Nov. 2	01 21	5 12	s.c.	-22	-47	+40	ms	2	5	7	208	1000	130
	Nov. 27	23 51	28 14	s.c.*	-4	+47*	-72	m	28	1,2	7	128	917	125
	Nov. 30	07 00		s.c.	+29	-237	-162	ms	30	7	7	180	1165	123
	Dec.		03 24										
	Dec. 26	11 00	29 13				m	27	6	7	142	(1270)	(940)
									28	5				
	1960													
	Jan. 10	07 00	11 18				m	10	6,7,8	6	161	1290	89

SELECTED GEOMAGNETIC AND SOLAR DATA

Kp, *Ci*, *Cp*, *Ap*, *K_F*, *Rz*, and Selected Days
December, 1960

Day ¹	Three-hour Range Indices Kp^2								Sum	Prel. ³ Ci	Cp^4	Ap^5	3-hr Range Indices K_{Pr}^6		Prov. ⁷ Rz	
	1	2	3	4	5	6	7	8					Values	Sum		
1 D	6+	8o	7-	5o	5+	6o	6-	6o	49o	1.8	1.9	93	5754	5445	39	74
2 D	4-	4o	4-	4o	5o	3+	3+	4-	31-	1.2	1.2	26	3334	4333	26	92
3 q	2+	4o	2o	3-	1+	2-	0+	1o	15+	0.3	0.5	9	2522	2101	15	101
4 Q	1-	0o	2o	2-	2+	2+	2o	1-	12-	0.3	0.2	6	0122	3110	10	96
5 q	1o	3o	1+	3o	3o	2+	3o	1+	18o	0.5	0.6	10	0313	3232	17	92
6	1o	3-	2+	4-	4+	4o	4+	4o	26+	1.2	1.1	21	2324	4323	23	98
7	2o	4-	2+	3-	4o	2-	5o	6-	26o	1.3	1.2	25	2323	3144	22	94
8	5-	5+	4-	4o	3-	2-	1o	2-	25-	0.9	1.1	22	5533	3211	23	91
9	2o	4+	3+	4-	4+	3+	4-	3-	27+	1.1	1.0	20	2433	3332	23	97
10 q	3-	3-	1+	2o	2o	2+	3+	2o	18+	0.6	0.5	10	2322	2220	15	103
11 Q	1o	1+	2+	2-	3-	3-	3-	3o	17+	0.4	0.5	9	1121	2322	14	102
12	3o	3o	2+	3o	3-	3o	4o	4+	25+	1.1	1.0	18	2223	2334	21	101
13 q	4-	4-	4-	3o	3+	1o	1o	1-	20o	0.6	0.8	14	2332	3001	14	92
14 Q	0o	1-	2o	3o	2+	2o	1-	1o	12-	0.3	0.3	6	0122	2201	10	101
15 D	3-	3o	3+	4o	6o	5+	5+	6o	36-	1.5	1.5	43	2334	4335	26	104
16 D	6-	6-	5o	5-	4-	2o	1+	2+	30+	1.2	1.3	33	5345	3112	24	99
17 Q	1-	2-	3o	2o	2-	1+	2o	2o	14+	0.2	0.4	7	0231	1122	12	87
18	3-	4o	5o	5-	3+	4o	4o	3-	30+	1.3	1.2	26	2344	3332	24	76
19	3+	3+	3+	4-	3-	3-	3-	4o	26-	0.9	0.9	17	3233	3223	21	70
20	4o	4o	3+	4o	3-	3-	5-	4o	29+	1.2	1.1	24	4434	3243	27	62
21	5o	3o	2+	2+	3+	4+	4o	5-	29o	1.2	1.2	25	5222	2334	23	56
22	4-	4o	4o	4-	3-	3o	3-	4-	27+	0.9	1.0	20	4443	2233	25	35
23	3+	3+	3o	3+	4o	2o	2o	2-	23-	0.7	0.8	14	3333	3111	18	35
24 q	2+	3-	1+	2+	4-	4-	3o	4o	23o	1.0	0.9	15	3212	3333	20	37
25 Q	4-	2o	2-	2o	2-	1-	3-	2-	16o	0.3	0.5	9	4222	1132	17	57
26	4o	4o	4o	3-	3+	2o	2-	2-	23+	0.8	0.9	16	4443	3222	24	48
27 D	2o	4o	5o	5o	5o	6o	6+	5o	38+	1.6	1.6	50	2454	4444	31	70
28	4o	3-	3+	5+	4o	4-	3+	2-	28o	1.1	1.1	23	3224	3321	20	86
29	4o	3+	4-	3+	2+	5-	3-	2o	26o	1.0	1.0	19	4233	2432	23	94
30	2o	4-	3+	2+	3-	4o	4o	3-	25-	0.9	0.9	17	1433	2342	22	103
31	2-	3+	4-	3o	3+	4-	3o	2o	24-	1.0	0.9	15	2333	3421	21	130
Means:										0.92	0.94	21			83.3	
No. of days:										31	31	31			31	

Notes:

¹ Five quiet days (Q), ten quiet days (Q or q), five disturbed days (D) selected by Committee on Characterization of Magnetic Disturbances, J. Veldkamp, Kon. Nederlandsch Meteorologisch Instituut, DeBilt Holland.

² Geomagnetic planetary three-hour-range indices *Kp* prepared by Committee on Characterization of Magnetic Disturbances, J. Bartels, Chairman, University, Göttingen, Germany.

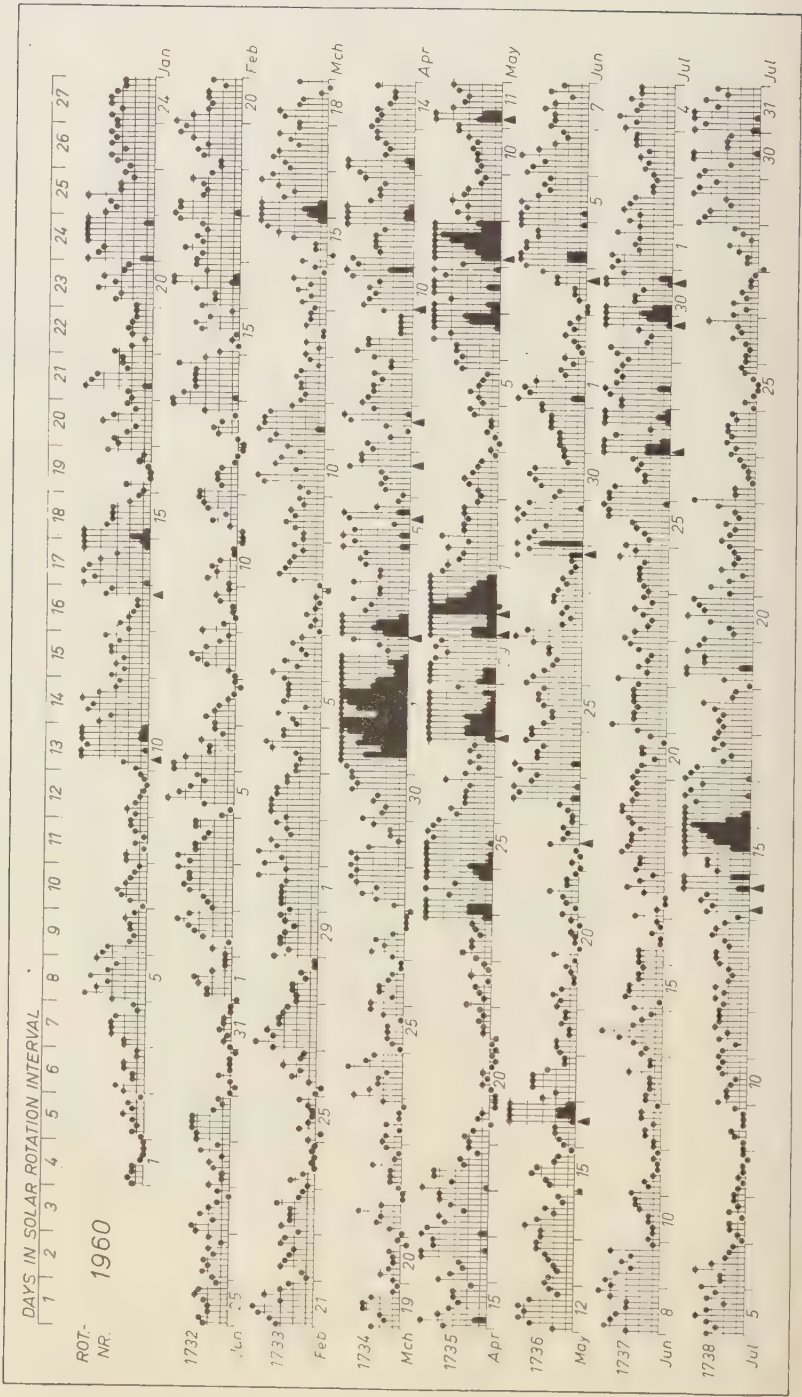
³ Preliminary magnetic character-figures, *Ci*, prepared by J. Veldkamp.

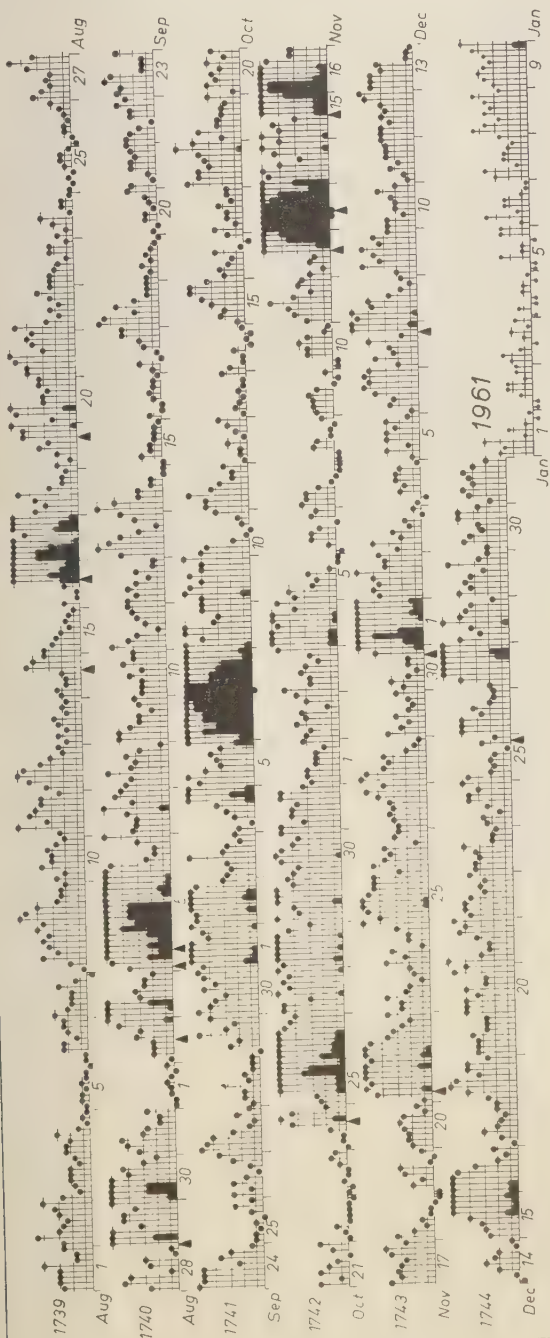
⁴ Magnetic character-figures, *Cp*, prepared by J. Bartels.

⁵ Average amplitudes *Ap* (unit 2γ), prepared by J. Bartels.

⁶ Fredericksburg three-hour-range indices *K* (*K*₉ = 500γ); scale-values of variometers in γ/mm: *D* = 2.7; (*H* = 2.5, *Z* = 3.2) prepared by Robert E. Gebhardt, Observer-in-Charge, Fredericksburg Magnetic Observatory, Corbin, Virginia.

⁷ Provisional sunspot-numbers (dependent on observations at Zurich Observatory and its stations at Locarno and Arosa) prepared by M. Waldmeier, Swiss Federal Observatory, Zurich, Switzerland.





PLANETARY MAGNETIC THREE-HOUR-RANGE INDICES K_p

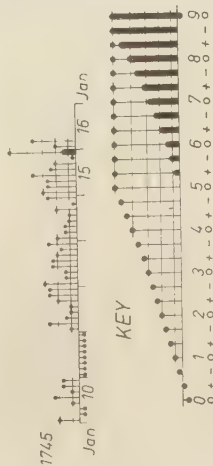
(and preliminary indices to 1961 January 16)

▲ = sudden commencement

1960

J. Bartels, Chairman

IAAGA Committee on Characterization of Magnetic Activity



Letters to the Editor

Short-Term Variations in Meson and Nucleon Component of Cosmic Rays

V. L. PATEL AND K. MAEDA

*University of Maryland
College Park, Maryland*

It has been shown that the relative decrease in nucleon intensity at two stations above the 'knee' is the same during several Forbush decreases [Fenton, McCracken, Rose, and Wilson, 1959]. Comparison of relative changes in nucleon intensity made at Ottawa and Hobart for eleven Forbush decreases during 1956-1958 found a ratio ~ 0.93 . McCracken [1959] calculated similar ratios for nucleon intensity changes at Hobart and Mawson; his average ratio was ~ 1.0 . However, the comparison of nucleon and meson intensity at Hobart gave ratios of 1.8-3.8, and 3.4-3.7 in two selected periods; the variation apparently is due to residual atmospheric effects in the meson intensity.

The purpose of this note to point out that a fairly constant ratio for decreases in nucleon and meson components can be calculated if the meson data are corrected by using radiosonde data. It has become clear that atmospheric effects are more predominant in the polar regions [Patel and Maeda, 1960] and that changes in the upper atmosphere do affect the sea-level meson intensity [Elliot, 1952; Sarabhai and Nerurkar, 1956].

TABLE 1. The Comparison of Decreases in Meson and Nucleon Components at Thule, Greenland

Only pressure correction applied in the meson intensity.

1958	$(\Delta I/I)_M' \%$ (Corrected for Pressure Only)	$(\Delta I/I)_N \%$	$(\Delta I/I)_N$ $(\Delta I/I)_M'$
Feb. 11	3.88	6.00	1.55
March 26	5.82	6.40	1.10
April 16	2.08	2.00	0.95
July 9	3.88	5.40	1.63
Oct. 22	2.77	3.80	1.37
Avg. = 1.32			

We have compared data from a meson telescope (standard cubical geometry) and from a neutron monitor, both located at Thule ($\lambda = 89^\circ \text{N}$). Pressure-corrected daily mean intensity during the IGY period was used, and the relative intensity changes $\Delta I/I$ were calculated by the same method used by Fenton and his co-workers. The results are shown in Table 1 for five events. The ratios vary considerably from event to event. The average value is 1.32; the dispersion is so high that the lowest value differs 28 per cent, and the highest value differs 23 per cent, from the average.

Next, the meson intensity was corrected for 300-mb height level using radiosonde data. The coefficient -8.3 per cent/km, calculated by linear correlation, was used throughout the period; and

$$\alpha = (\Delta I/I)_N / (\Delta I/I)_M$$

was calculated again. The value of the temperature coefficient is fairly large as compared with those corresponding to middle and low latitude cosmic ray stations. The explanation of this large value of coefficient is given in another paper (Maeda and Patel, 1960). The results are shown in Table 2. The average value is $\alpha = 1.22$. The error in daily mean of meson intensity if 0.09 per cent and in nucleon intensity 0.14 per cent; further, 0.1 per cent is added to both to account for inaccuracy in atmospheric correction and random error. The simple calculation gives now error in each ratio equal to ± 0.31 ; hence an error of ± 0.13 in the average α . The dispersion from the average is very low now. The lowest value of α differs by 3 per cent and the highest value by 5 per cent from the average.

It is interesting that Ehmert [1958] had compared total meson intensity (recorded by double coincidence) and neutron monitor data at Weissenau; after applying a correction for the

TABLE 2. Comparison of Meson Intensity and Nucleon Intensity Decreases at Thule, Greenland

Meson intensity corrected for pressure and 300-mb level height.

1958	$(\Delta I/I)_M\%$ (Corrected for pressure and 300 mb H)	$(\Delta I/I)_N$	$(\Delta I/I)_N$ $(\Delta I/I)_M$
Feb. 11	5.10	6.00	1.18 ± 0.31
March 26	5.40	6.40	1.19 ± 0.31
April 16	1.60	2.00	1.25 ± 0.31
July 9	4.50	5.40	1.20 ± 0.31
Oct. 22	3.20	3.80	1.28 ± 0.31
Avg. = 1.22 ± 0.14			

300-mb level, he obtained ratios from 1.1 to 2.5. However, our result at Thule from the hard component and the nucleon component gives a fairly constant ratio of 1.22 after correction is applied for 300-mb level in the meson intensity.

It is noteworthy from our analysis that at sea level the neutron intensity changes are only 20 per cent greater than meson decreases during Forbush decreases.

This determination of α may be used to remove truly world-wide changes in either component in order to evaluate atmospheric effects. The ratio α at a particular place on the earth gives relative changes of world-wide character in both components. These changes now can be eliminated from meson component by calculating $(M - N/\alpha)$. The quantity $(M - N/\alpha)$ gives the

variation in meson intensity contributed mostly by residual atmospheric effects. From this in turn more accurate correction for atmospheric effects can be estimated.

Acknowledgments. We are grateful to Professor S. F. Singer for suggesting this investigation. We are thankful to the Director, National Weather Bureau, Asheville, North Carolina, for supplying radiosonde data, and to Dr. M. A. Pomerantz for supplying neutron monitor data from Thule, Greenland. Dr. M. J. Swetnick participated in the early phases of the Thule observations. This work was supported in part by National Science Foundation Grant G-8377.

REFERENCES

- Ehmert, A., Variations of nuclear and mesonic component, *Nuovo cimento Suppl.*, 8, 318-325, 1958.
 Elliot, H., *Progress in Cosmic Ray Physics*, vol. 1, North Holland Publishing Company, Amsterdam, 452-514, 1952.
 Fenton, A. G., K. G. McCracken, D. C. Rose, and B. G. Wilson, Transit decreases in cosmic ray intensity during the period October 1956 to January 1958, *Can. J. Phys.*, 37, 569-578, 1959.
 Maeda, K., and V. L. Patel; Seasonal variation of cosmic-ray intensity in polar regions, *J. Geophys. Research*, 66, in press 1961.
 McCracken, K. G., Variation in cosmic ray rigidity spectrum, *Phys. Rev.*, 113, 343-348, 1959.
 Patel, V., and K. Maeda, Seasonal variation of cosmic rays in polar regions, *Bull. Am. Phys. Soc.*, 5, 23, 1960.
 Sarabhai, V., and N. W. Nerurkar, Time variation of primary cosmic rays, *Ann. Rev. Nucl. Sci.*, 6, Annual Reviews, Inc., Stanford, 6-10, 1956.

(Received November 28, 1960; revised January 12, 1961.)

A Note on Solar Flare Cosmic Rays

K. MAEDA AND V. L. PATEL

University of Maryland, College Park, Maryland

Recently, several workers have investigated the effects of solar flares upon the cosmic-ray intensity by means of Chree's analysis or the method of superposition. An apparent contradiction exists among their conclusions. This brief note is intended to point out a possible interpretation for these discrepancies in terms of the selection of data used in their analyses.

According to the Chree analysis done by Towle and Lockwood [1959], and Ghielmetti, Anderson, Cardoso, Manzano, Roederer, and Santochi [1960], no significant increase in cosmic-ray intensity with solar flares is detected during the period from 1956 to 1957, and from July 1957 to August 1958; respectively. On the other hand, Kolomeets [1959] found significant increases at four stations in Europe for the period of several months in 1957.

In all three analyses, the intensity of cosmic-ray neutrons corrected for barometric pressure is used; the only essential difference between the former and the latter is the consideration on impact zones; namely, Towle-Lockwood and Ghielmetti and his co-workers have restricted their analyses to events in the impact zones, but Kolomeets did not.

To investigate this point, Patel [1960] has made a Chree analysis of cosmic-ray data obtained by the University of Maryland Cubical

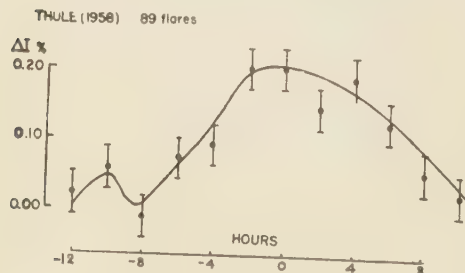


Fig. 1. Variation of cosmic-ray intensity following solar flares at Thule, Greenland (geomag. lat. 89°N).

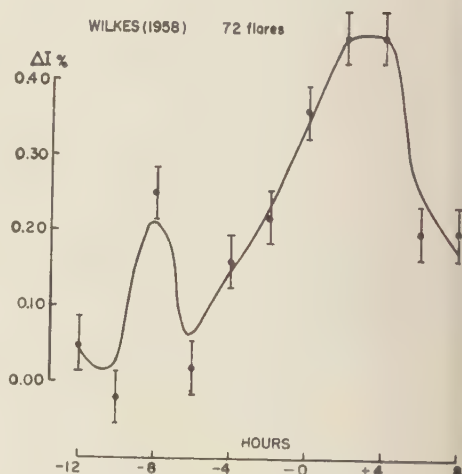


Fig. 2. Variation of cosmic-ray intensity following solar flares at Wilkes, Antarctic (geomag. lat. 76°S).

Meson Telescope during the IGY at the polar stations, one at Thule, Greenland (geomagnetic latitude 89°N) and the other at Wilkes, Antarctica (geomagnetic latitude 76°S).

Figures 1 and 2 show the results for Thule obtained by 89 events of solar flare, and for Wilkes with 72 events, respectively. In both instances meson data are corrected for barometric effect and the zero hour in abscissa indicates the time of onset of visible solar flares of importance larger than 2. Although there are no impact zones in the polar regions for cosmic-ray particles, the enhancement of cosmic-ray intensity following the solar flares seems genuine. Notice, however, that during the active period of the sun, the occurrence of a flare is quite frequent; therefore, the tendency for a cosmic-ray increase earlier than 2 hours before flare-onset seen in these figures is due not only to the uncertainty of the bi-hourly cosmic-ray data used in the analysis but also to the tails of preceding events in the

chosen epochs. This also is seen in Kolomeets' analysis as a small dip around 4 hours before the zero epoch time. Moreover, contrary to Kolomeets' analysis, apparent increase in cosmic-ray intensity due to the diurnal variation can be excluded in our analyses, because no significant diurnal variation of cosmic-ray intensity exists in polar regions.

The above analysis indicates, therefore, that (1) all solar flares produce cosmic-ray particles to some extent as suggested by *Singer* [1958], and (2) owing to the disturbed nature of the magnetic field in the solar system during the period of solar maximum, many of the emitted solar cosmic rays do not follow the simple orbits, on which the calculation of impact zones is based [*Ehmert*, 1948; *Schlüter*, 1951; *Firor*, 1954; *Jory*, 1956; *Lüst*, 1957].

Finally, it should be noticed that during the period of weak solar activity, most solar cosmic rays might be able to follow a Störmer type orbit from the vicinity of the sun to the top of the earth's atmosphere, giving noticeable increases in the impact zones as shown by *Firor's* [1954] analysis.

Acknowledgments. We wish to express our gratitude to Professor S. F. Singer who encouraged us to work on the present analyses. We are grateful to Messrs. F. J. Pavlica, Jr., J. B. MacWherter

and J. A. Jones for maintaining cosmic-ray equipment at Thule and at Wilkes. Dr. M. J. Swetnick participated in the early phases of the work. This work is supported by National Science Foundation Grant, G-8377.

REFERENCES

- Ehmert, A., Ultrastrahlung von der Sonne, *Naturforscher*, *3a*, 264-285, 1948.
 Firor, J., Cosmic radiation intensity—Time variations and their origin IV, *Phys. Rev.*, *94*, 1017-1728, 1954.
 Ghielmetti, H., J. C. Anderson, J. M. Cardoso, J. R. Manzano, J. G. Roederer, and O. R. Santochi, Solar flare effects on cosmic ray intensity, *Nuovo cimento*, *15*, 87-98, 1960.
 Jory, F. S., Selected cosmic ray orbits in the earth's magnetic field, *Phys. Rev.*, *103*, 1068-1075, 1956.
 Kolomeets, E. V., Small effects of solar flares and energy spectrum of primary variation of cosmic rays, *Soviet Phys., JETP*, *36* (9), 960-961, 1959.
 Lüst, R., Impact zones for solar particles, *Phys. Rev.*, *105*, 1827-1839, 1957.
 Patel, V. L., Time variations of cosmic rays in the polar regions, Master's Thesis, University of Maryland, 1960.
 Singer, S. F., A model for solar flare increases of cosmic rays, *Nuovo cimento*, Suppl., *8*, 197-201, 1958.
 Towle, L. C., and J. A. Lockwood, Cosmic ray increase associated with solar flares, *Phys. Rev.*, *113*, 641-644, 1959.

(Received November 28, 1960; revised January 12, 1961.)

Solar-Stream Distortion of the Geomagnetic Field and Polar Electrojets

J. W. KERN

*Department of Planetary Sciences, The RAND Corporation
Santa Monica, California*

Distortion of the geomagnetic field by an ionized solar stream has often been thought to be related to the occurrence of polar electrojet current systems and aurora [Chapman and Bartels, 1951; and others]. Such distortion has been considered by Chapman and Ferraro [1931], Chapman [1960], and Ferraro [1960], Martyn [1951], and more recently by Piddington [1959, 1960], who considers the effects of the plasma content of the earth's exosphere, and suggests the possibility of a geomagnetic tail due to interaction with a solar stream. Distension of the geomagnetic field due to trapped particle currents [Dessler and Parker, 1959; Akasofu, 1960] may also contribute to such distortion. Chapman [1950] formulated an experiment to clarify the physical effects that involved motion of an ionized gas past a magnet, but scaling difficulties prevented such an approach. Johnson [1960] has considered the over-all problems, introducing recent developments in plasma physics. If simultaneous day-side compression and night-side stretching of geomagnetic field lines is considered, the night-side stretching may be limited by the dynamic pressure of the solar stream [Martyn, 1951].

It is the purpose of this letter to point out that magnetic field gradients perpendicular to meridional planes may be introduced by such distortion with the creation of polar electrojets. It is claimed that longitudinal magnetic field gradients will lead to drift separation of geomagnetically trapped protons and electrons normal to field lines and in meridional planes, and thus to a suitable latitudinal charge separation in trapped radiation incident at auroral latitudes. Penetration of such polarized trapped radiation to the *E* region of the atmosphere is supposed to lead to electrojet current systems of the form observed. The work here continues the study of the concepts recently considered by

Vestine [1960] and by Chamberlain, Kern, and Vestine [1960].

Initial irregularities in the plasma number density are necessary in order for drift separation due to longitudinal magnetic field gradients to produce polarization in an initially neutral plasma. In the absence of drift separation, any polarization electric fields will decay rapidly. It would appear, however, that where such irregularities in number density and drift separation occur together, a mechanism may exist for continuous supply of polarized trapped radiation to auroral regions.

If polar electrojet current systems are ascribed to Hall conduction in the *E* region driven by meridional electric fields [Baker and Martyn, 1953; Chamberlain, Kern and Vestine, 1960; Akasofu, private communication], eastward-directed electrojets are then associated with poleward-directed electric fields and westward-directed electrojets with equatorward-directed electric fields. The direction of drift of a proton due to longitudinal magnetic field gradients is given by $v_{dp} = (\mu_p/eB^2) \mathbf{B} \times \nabla_\phi B$, where μ_p is the proton magnetic moment, e the electronic charge (here considered positive), \mathbf{B} the magnetic field, and $\nabla_\phi B$ the longitudinal component of the magnetic field gradient. A similar expression applies for an electron, but the direction of drift is opposite that of a proton, because of the opposite sign of the electric charge. Thus the total drift separation velocity is given by $v_d = [(\mu_p + \mu_e)/eB^2] \mathbf{B} \times \nabla_\phi B$. The drift directions produced by eastward-directed longitudinal magnetic field gradients are such as to produce poleward electric polarization fields, while westward-directed magnetic gradients produce equatorial polarization fields in the trapped radiation incident in auroral regions. Eastward-directed electrojets may therefore be associated with eastward-directed longitudinal

magnetic field gradients, and westward-directed electrojets with westward-directed gradients.

The extent in longitude of such meridional electric fields would correspond to the extent of the longitudinal magnetic field gradients where regularities in plasma density exist to allow the development of polarization. Electrojets could develop where sufficient trapped radiation, polarized by the above mechanism, penetrates the ionosphere, perhaps accompanying mirror-point lowering due to geomagnetic field distortion [Vestine, 1960] or diamagnetic effects of charged-sheet geometries [Kern and Vestine, 1961]. Continuous supply of particles to conduction currents in the ionosphere would lead to smoothing-out of the original plasma density irregularities and hence a diminution of current supplied as a function of time.

The process described above requires that the direct conduction current associated with the full current electrojets be supplied by continuous drift separation of trapped particles. This current is equal therefore to the latitudinal drift charge transport $nev_d r l$ per radian of longitude in the auroral zone, where n is the plasma number density, $e = 1.6 \times 10^{-20}$ emu is the electronic charge, v_d is the drift separation velocity for particles of unlike sign, $r = 5 \times 10^9$ cm is the average radius of the region of drift separation conjugate to auroral latitudes, and $l = 2 \times 10^9$ cm is the approximate extent of the drift region along lines of force. The total direct conduction current per radian of longitude in the auroral region associated with a typical electrojet current can be estimated. If Baker and Martyn's [1953] value for the height-integrated direct conductivity of 10^{-8} emu and an estimated electric field of 10^4 emu are used, and if the radius of the auroral zone is taken as 2.5×10^8 cm, this current turns out to be 2.5×10^4 emu per radian of longitude in the auroral zone. Equating this to the drift current given above, and taking $n = 10^3/\text{cm}^3$, permits calculation of the drift separation velocity required to maintain such a current as $v_d = 150$ cm/sec.

We can apply the equation developed earlier for the drift separation velocity to obtain an estimate of the geomagnetic field distortion required to produce such drift velocities. For 10-keV particles mirroring in the auroral region, the magnetic moment $\mu = 2 \times 10^{-8}$ emu, and hence the magnitude of the longitudinal com-

ponent of the magnetic field gradient can be written $|\nabla_\phi B| = (l/r) dB/d\phi$, we find that $dB/d\phi = (Ber/2\mu)v_d$. With $B = 10^{-3}$ gauss at 7 to 8 earth radii, $e = 1.6 \times 10^{-20}$ emu, $r = 5 \times 10^9$ cm, and $v_d = 150$ cm/sec, we obtain $dB/d\phi = 3 \times 10^{-4}$ gauss/radian. This amount of distortion is quite modest and would be considered applicable to times of magnetic bays.

If a dynamic pressure model is assumed for the form of the geomagnetic field distortion in an ionized solar stream moving nearly radially from the sun, the magnetic field gradients may be eastward before magnetic midnight and westward after magnetic midnight. The current systems and meridional electric fields associated with such gradients are, according to the above theory, such as to account for the magnetic observations of Heppner [1954], for College, Alaska. Observed longitudinal drift of auroral structures are accounted for in terms of $\mathbf{E} \times \mathbf{B}$ motor drift, westward drift being predicted before midnight and eastward after midnight [Nichols, 1957; Kim and Currie, 1958]. Further, the observed southward drift of auroral structures [Leadabrand, Presnell, Berg, and Dyce, 1959; and others] can be identified with the predicted southward latitude drift of incident electron sheets near magnetic midnight. The theory seems also to predict incident proton drift near magnetic midnight in the opposite direction; i.e., northward [Rees and Reid, 1960], with cross-over of proton and electron drift at solar-stream midnight.

The drift separation mechanism outlined above has the additional feature of separating particles of opposite sign into sheet geometries and maintaining such charge-density distributions in the trapped radiation. This allows for charge sheets in the geomagnetic field [Kern and Vestine, 1961]. If the region of geomagnetically trapped particles has a fairly sharp outer boundary adjacent to the northern auroral zone, the present theory suggests that before magnetic midnight there will be formation of an electron sheet boundary. After midnight, the formation of a proton sheet on the boundary would be expected. Diamagnetic effects of such sheets of charges would lead to a lowering of mirror points and, hence, to visible aurora. The fixed spatial relation of the auroral pattern with respect to the sun observed by Davis [1960] would, of course, be predicted

by this theory. Concentric patterning of auroral sheets, reported by Davis, would be accounted for on the basis of the diamagnetic sheet mechanism [Kern and Vestine, 1961]. Since the number of electrons and protons per unit volume are unequal, motor drift in the direction $\mathbf{E} \times \mathbf{B}$, will arise. On either side of the sheet the current directions are reversed. Outward forces between these current sheets will result, leading to expansion of the original sheet into a loop. If the incident particles have a discrete spectrum of velocities along lines of force, the drift mechanism may serve to analyze the charge supply that was homogeneous in the outer Van Allen region into a series of concentric loops, each formed by particles of like energy [McIlwain, 1960].

Recurrence of electrojets and aurora at about the same longitude on successive nights can be expected if low-energy particles, such as the 6-keV auroral electrons observed by McIlwain, predominate in the trapped radiation at conjugate auroral points, since the earth rotates within the distorted field pattern. Such particles drift very slowly in longitude with respect to the earth's surface, hence would effectively rotate with the geomagnetic field within the distorted cavity produced by the solar stream. Thus an increased number density of low-energy particles entering the high atmosphere at conjugate points in northern and southern auroral regions, at 24-hour intervals, would exhibit a polarization determined by the particle drift in the distorted geomagnetic field. It seems likely that scattering to lower mirror points [Welch and Whitaker, 1959] would replenish the particle supply for aurora and electrojets following depletion of low mirror-point particles by a given display.

REFERENCES

- Akasofu, Syun-Ichi, The ring current and the outer atmosphere, *J. Geophys. Research*, **65**, 535-543, 1960.
- Baker, W. G., and D. F. Martyn, Electric currents in the ionosphere, *Phil. Trans. Roy. Soc. London*, **A**, **246**, 281-320, 1953.
- Chamberlain, J. W., J. Kern, and E. H. Vestine, Some consequences of local acceleration of auroral primaries, *J. Geophys. Research*, **65**, 2535-2537, 1960.
- Chapman, S., Idealized problems of plasma dynamics relating to geomagnetic storms, *Rev. Modern Phys.*, **32**, 919, 1960.
- Chapman, S., Notes on auroral and magnetic storms, *J. Atmospheric and Terrest. Phys.*, **1**, 189-199, 1950.
- Chapman, S., and J. Bartels, *Geomagnetism*, II Oxford University Press, 1951.
- Chapman, S., and V. C. A. Ferraro, A new theory of magnetic storms, Part 1, The initial phase *Terrestrial Magnetism and Atmospheric Elec.*, **36**, 77-97, 1931.
- Davis, T. Neil, The morphology of the polar aurora *J. Geophys. Research*, **65**, 3497-3500, 1960.
- Dessler, A. J., and E. N. Parker, Hydromagnetic theory of geomagnetic storms, *J. Geophys. Research*, **64**, 2239-2252, 1959.
- Ferraro, V. C. A., Theory of sudden commencements and of the first phase of a magnetic storm *Rev. Modern Phys.*, **32**, 934, 1960.
- Heppner, J. P., Time sequences and spatial relations in auroral activity during magnetic bays at College, Alaska, *J. Geophys. Research*, **59**, 329-338, 1954.
- Johnson, F. S., The gross character of the geomagnetic field in the solar wind, *J. Geophys. Research*, **65**, 3049-3051, 1960.
- Kern, J. W., and E. H. Vestine, Theory of auroral morphology, *J. Geophys. Research*, **66**, 713-723, 1961.
- Kim, J. S., and B. W. Currie, Horizontal movements of aurora, *Can. J. Phys.*, **36**, 160-170, 1958.
- Leadabrand, R. L., R. I. Presnell, M. R. Berg, and R. B. Dyce, Doppler investigations of radar aurora at 400 Mc/s, *J. Geophys. Research*, **64**, 1197-1203, 1959.
- Martyn, D. F., The theory of magnetic storms and auroras, *Nature*, **167**, 92-98, 1951.
- McIlwain, C. E., Direct measurements of particle producing visible auroras, *J. Geophys. Research*, **65**, 2727-2747, 1960.
- Nichols, B., Drift motions of auroral ionization, *J. Atmospheric and Terrest. Phys.*, **11**, 292-293, 1957.
- Piddington, J. H., The transmission of geomagnetic disturbances through the atmosphere and interplanetary space, *Geophys. J.*, **2**, 173-189, 1959.
- Piddington, J. H., Geomagnetic storm theory, *J. Geophys. Research*, **65**, 93-106, 1960.
- Rees, M. H., and G. C. Reid, Spectroscopic development of the aurora, *Planetary Space Sci* in press, 1960.
- Vestine, E. H., Polar auroral, geomagnetic, and ionospheric disturbances, *J. Geophys. Research*, **65**, 360-362, 1960.
- Welch, J. A., and W. A. Whitaker, Theory of geomagnetically trapped electrons from an artificial source, *J. Geophys. Research*, **64**, 909-922, 1959.

(Manuscript received December 8, 1960.)

Some Results of Direct Probing in the Ionosphere

W. PFISTER, J. C. ULWICK, and R. P. VANCOUR

*Air Force Cambridge Research Laboratories
L. G. Hanscom Field, Bedford, Massachusetts*

Two different techniques of direct probing in the ionosphere were combined in a rocket experiment on August 28, 1959. The two probes were the antenna impedance probe [Haycock and Baker, 1960] and the retarding potential probe [Tinteregger, 1960].

In order to minimize the effect of the ion sheath around the rocket and the antenna, a symmetrical dipole arrangement has been used consisting of a pair of whips that are erected after burnout of the rocket motor. Loading coils wound around the whips outside the rocket body ensured a minimum RF voltage at the feed points and made the free space reactance inductive. The RF voltage on the antenna was kept less than 1 volt rms. Additionally, a d-c voltage sweeping from 0 to 4 volts in a period of 2 seconds was applied between the rocket body and the antenna. This voltage excursion was sufficient to force the antenna through the undisturbed space potential. Two completely independent measurements at 3.0 Mc/s and 7.2 Mc/s were carried out simultaneously with the impedance going through resonance in both cases.

For the first analysis, the effect of the earth's magnetic field was neglected and the collision frequency was assumed to be known. This allowed us to compute the electron density using only the reactive component according to the following formula:

$$N = \frac{f_m^2 \cdot 10^6}{80.6} \frac{2\omega C_0 \Delta X + 1 - \sqrt{1 - (\nu/\omega)^2 \cdot 4\omega C_0 \Delta X (\omega C_0 \Delta X + 1)}}{2(\omega C_0 \Delta X + 1)}$$

where

ΔX is the change in the reactive component from the free-space value.

C_0 is the free-space capacitance of the antenna.
 ω is the operating frequency.

Figures 1 and 2 show the electron density results obtained from the 3.0 and the 7.2 Mc/s

data. Antenna erection was between 59 and 65 km. A complete mismatch occurred at 87 km for the 3.0-Mc/s frequency and at 92 km for the 7.2-Mc/s frequency. The general shape of the electron density curves appears reasonable, although the amplitude is slightly higher than expected.

No data for descent are available for comparison since telemetry ceased functioning below 140 km. However, due to a phenomenon still to be investigated, measurements could be made at 7.2 Mc/s in the altitude range from 200 to 140 km during descent. The results are interpreted as relative electron densities which appear to require a correction by a factor of 6 in order to be comparable with the ion probe results and the true height curve (Fig. 3). The latter has been computed from the ionogram provided by the White Sands station but is subject to some uncertainties because there was not any normal *E*-layer trace visible on the record. The electron density fluctuations appear larger in amplitude than at the lower height. This is explained partly by the effect of the magnetic field on the impedance at the higher altitude. This will be studied in more detail when the rocket aspect data have been reduced.

Although there is an ion sheath around the antenna which is altered by a d-c sweep voltage between rocket body and antenna with a cor-

responding measurable Langmuir current, not the slightest effect on the impedance was detected either in the *D* region or in the lower *F* region. This indicates that with an antenna potential of about ± 2 volts relative to space, and an antenna of measurable radiation efficiency, the effect of the ion sheath on an impedance probe measurement is negligible. This important result

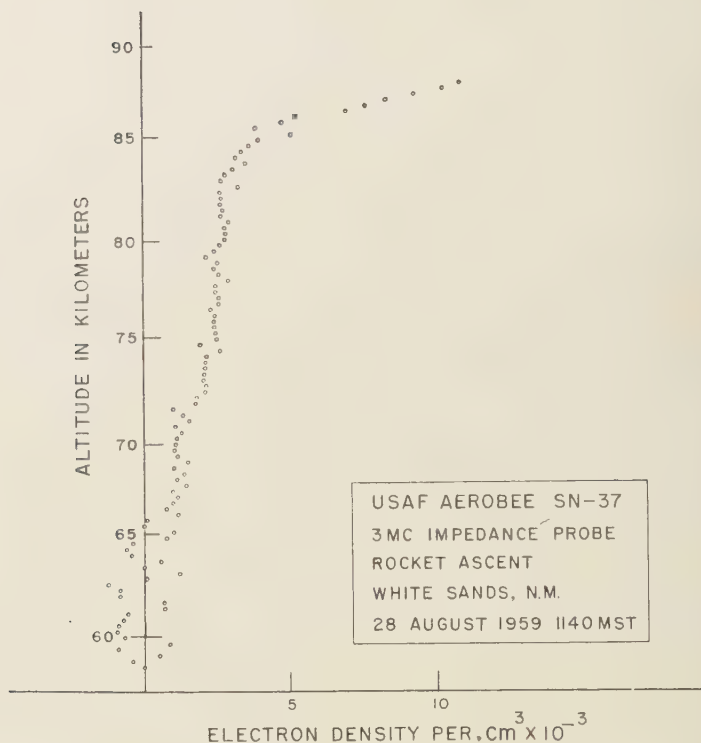


Fig. 1. Electron density impedance probe measurements.

was confirmed by a later similar flight in May 1960.

Since the rocket was very stable during the flight and pointed vertically upward during ascent and descent, the analysis of Hinteregger's ion probe data is simplified appreciably. Assuming a single ion species with a Maxwellian velocity distribution, and considering that the plane of the electrodes (grids and target plate) was nearly perpendicular to the velocity vector of the vehicle, we can use the simple equation

$$i = eA \int_{-V}^{\infty} (v + V) \frac{n}{\alpha \sqrt{\pi}} \exp \left[-\left(\frac{v}{\alpha} \right)^2 \right] dv$$

for the unretarded current of positive ions collected by the target plate during ascent. Here α represents the most probable velocity of the positive ions, n is their density, V is the vertical component of the rocket velocity, A is the 'effective' area of collection (assumed to be equal to the actual detector aperture in this analysis),

and the integration variable v refers to the component of ion velocity perpendicular to the aperture plane.

Using the $+$ sign for ascent (i_A) and the $-$ sign for descent (i_D), we obtain

$$i_{A,D} = \frac{enAV}{2} \left\{ \pm 1 + \operatorname{erf} \left(\frac{V}{\alpha} \right) + \frac{1}{\alpha} \frac{1}{\sqrt{\pi}} \exp \left[-\left(\frac{V}{\alpha} \right)^2 \right] \right\}$$

By the difference $i_A - i_D$, it is possible to compute the positive ion density n .

Since the ratio of the currents for ascent and descent depends only on V/α , and since V is known, the most probable value of thermal ion velocity α can be computed. Once α is known, the corresponding temperature can be determined if a value is assumed for the ion mass.

The results of such an analysis are shown in

Figure 4. Below 170 km, no descent data were available. Thus, no estimate of α or the temperatures could be given. However, the positive ion density below 170 km could be determined from the ascent data since the simple relation

$i = enAV$ is a good approximation for these altitudes where $V \gg \alpha$.

The dotted triangles in Figure 4 mark ion density values obtained by adopting the temperatures of the 1959 ARDC Model Atmosphere

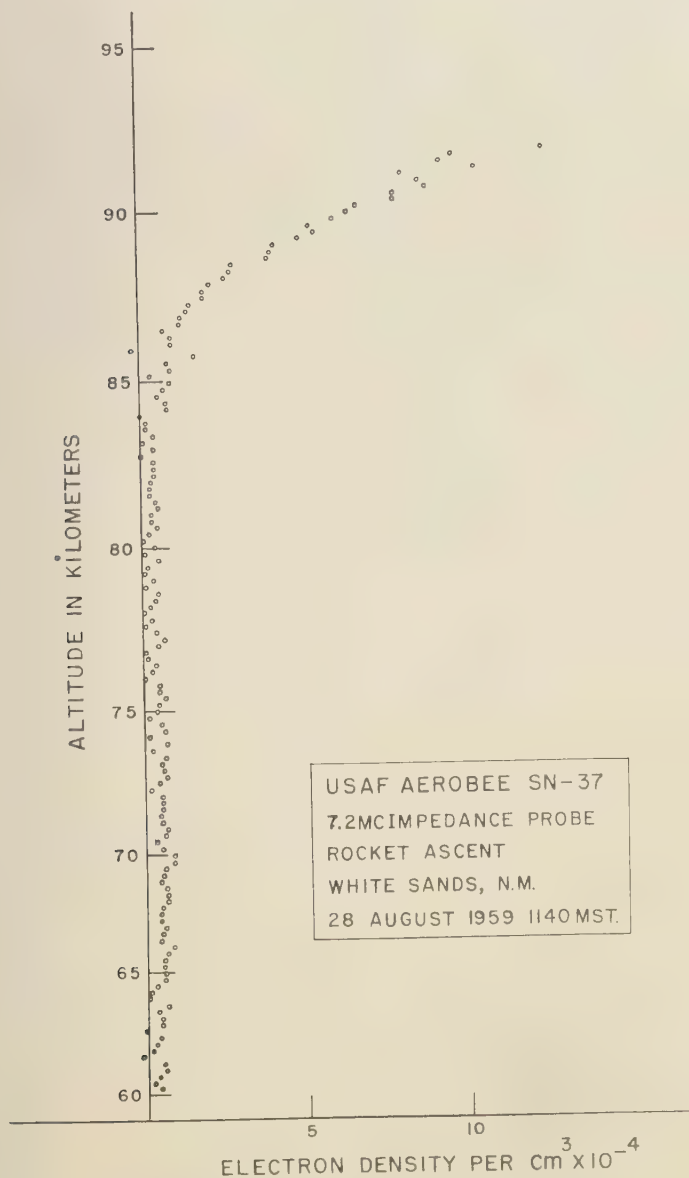


Fig. 2. Electron density impedance probe measurements.



Fig. 3. Electron density impedance probe measurements.

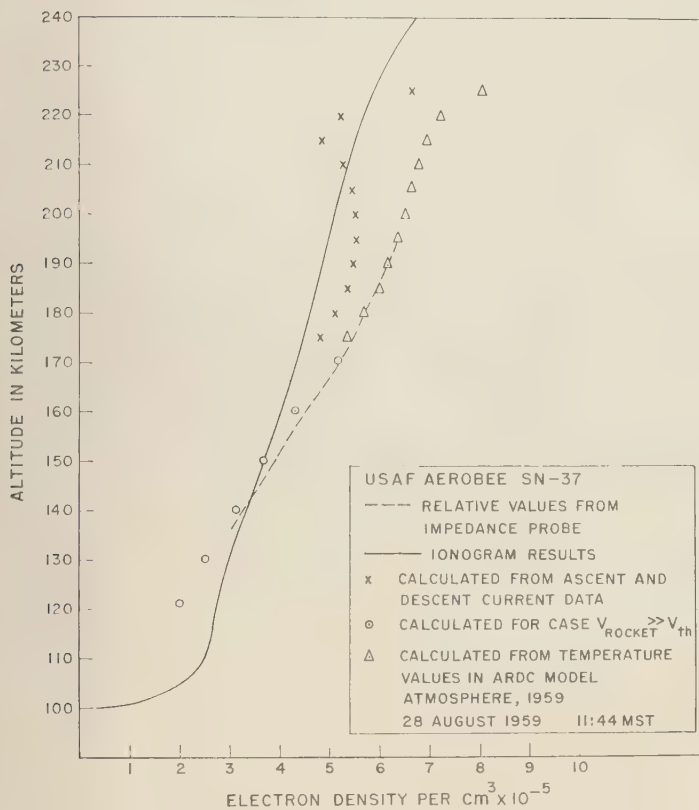


Fig. 4. Ion density measurements.

and assuming the ion mass to be 16 amu (O^+). The lack of agreement with the ion density marked by crosses (not implying the assumption of the model temperature) could be remedied by assuming much higher temperatures. More reasonable, however, is the introduction of a correction which takes into account the deviation from a plane geometry. This should bring the lower curve up (marked by crosses) and the upper curve down (marked by triangles). A more refined analysis is necessary to make a good quantitative evaluation of the ion concentration.

REFERENCES

- Haycock, O. C., and K. Baker, Measurement of antenna impedance in the ionosphere, *Electronics*, submitted November 1960.
- Hinteregger, H., Combined retarding potential analysis of photoelectrons and environmental charged particles up to 234 km, *Proceedings First International Space Science Symposium*, Nice, France, North-Holland Publishing Co., 1960.

(Received December 10, 1960; revised February 7, 1961.)

Air Density Variations in the Mesosphere, and the Winter Anomaly in Ionospheric Absorption

J. MAWDSLEY

Dominion Physical Laboratory, Lower Hutt, New Zealand

A recent paper by Jones, Peterson, Schaeffer, and Schulte [1959] reporting large mesospheric air density variations at high latitudes in winter may be of considerable interest regarding the anomalous winter daytime absorption of medium- and high-frequency radio waves. A plot of noon absorption of vertically reflected 2 Mc/s radio waves at Slough from year to year shows that the absorption in winter is usually much greater than would be expected if it were simply related to the zenith angle of the sun [Whitehead, 1957]. Noon absorption has been measured over long periods at low as well as at high latitudes, but the effect has been reported only in the event of the latter. It is not present, for instance, in the absorption measured at Singapore and published in the Slough Bulletin. Although the phenomenon has been known for many years, its physical cause remains in doubt.

The purpose of this note is to point out that in any future series of mesospheric air density measurements, particularly at higher latitudes, it would be well worthwhile to have a simultaneous indication of ionospheric absorption at the same place. A simple way of monitoring ionospheric absorption has been suggested by recording the sky-wave field-strength of a suitable medium- or high-frequency transmitter [Beynon and Davies, 1954]. Alternatively, values of f_{\min} from a nearby ionosonde maintained at constant sensitivity and gain might be adequate in this respect, and might permit a comparison on the basis of past experiments.

The extra absorption of E -reflected echoes in winter at 2 Mc/s is of the nondeviative kind [Whitehead, 1957]. That it is thus likely to be due to extra ionization near the level of the mesosphere is supported by the associated appearance, at lower frequencies, of echoes totally reflected from the same level [Gnanalingam and Weekes, 1955]. To account for the extra absorption on some days, the electron density must be several times the normal value. (Collision frequency changes accompanying the meso-

spheric air density changes reported by Jones, Peterson, Schaeffer, and Schulte [1959] are unlikely to be important in this respect.) Normal daytime ionization at this level is due chiefly to the action of solar Lyman- α radiation on nitric oxide [Nicolet and Aikin, 1960]. In principle extra ionization may come about through (a) an increase in the rate of ion production; (b) a decrease in the ionization loss rate, or (c) a movement of ionization from other levels.

An explanation involving movement of ionization meets formidable difficulties since it requires movement from higher levels where no ionization decrease is readily observed [Beynon and Davies, 1955] to lower levels where the collision frequency is higher, but where recombination proceeds more rapidly. Neither does the second suggestion appear to offer a satisfactory explanation. Although the actual processes of electron loss in the mesosphere may be open to question, the loss rate is fairly high. The disappearance of the extra absorption when the sun sets suggests that the extra ionization is removed by an efficient process. Since it is unlikely that, for some reason, the loss rate is low only on certain days in winter, it seems preferable to adopt the view that the extra ionization is not due to this cause.

Can we then, ascribe the increase to an increase in the rate of ion production? It is a well-known result of the Chapman theory of ion production that the maximum rate of ion production (for a particular value of solar zenith angle) depends only on the incident photon flux and the scale height of the constituent being ionized. In particular, the density of the ionizable constituent does not affect the maximum rate of ion production provided that its scale height is the same. What is required is a change in the incident photon flux. Thus it seems that the last explanation is untenable since winter days with extra absorption, although they tend to occur in groups, are not associated with solar activity [Appleton and Piggot, 1954]. However

solar Lyman- α radiation is absorbed chiefly in exciting molecular oxygen [Byram, Chubb, and Friedman, 1954]. Only a small fraction of it is absorbed in ionizing nitric oxide. Increases in the abundance of nitric oxide relative to molecular oxygen would have the same effect in ionizing nitric oxide as increases in the incident photon flux. Indeed, according to Nicolet [1955] atmospheric motions will affect the production and downward transport of nitric oxide, and increases in electron density at the mesospheric level are then to be expected even when the sun is quiet.

There appears then, to be a strong suggestion that the large air-density increase observed in winter at Fort Churchill by Jones and his co-workers may reveal vertical air movements that bring about the much larger enhancement in nitric oxide density suggested by Nicolet and give rise, in turn, to increased daytime radio wave absorption. Since this note was first prepared, evidence has appeared that air densities and movements may be linked in this way [Stroud, Nordberg, Bandeen, Bartman, and Titus, 1960]. It is, moreover, significant that the anomalous winter absorption appears also to be confined to higher latitudes, and that it is present (somewhat sporadically) only in winter. It is also interesting to note that the twilight enhancements of oxygen airglow reported recently by Megill, Jamnick, and Cruz [1960] follow a similar pattern. A one-to-one correspondence between the radio wave absorption and air density phenomena is not to be expected if the nitric oxide enhancement persists after the normal air density recovers.

In conclusion, a paper by Eyfrig [1953] is worthy of note. He showed that winter absorption anomaly observations correlate only over much smaller distances than would be expected if they are due to solar activity, and reported an instance where a group of days of high absorption was preceded by an extremely rapid temperature rise in the stratosphere. Jones and his co-workers report what appears to be a similar temperature rise associated with the air density fluctuations already mentioned. Simultaneous observations of the type suggested should thus help establish the cause of the anomalous winter absorption.

Additional lower altitude observations, such as those discussed recently by Allington, Boville, and Hare, [1960], of large disturbances confined

to winter and relatively high latitudes, lead one to wonder if the effect discussed above is an upper atmosphere manifestation of a more general meteorological feature. If this is so, then an association is indicated between meteorological behavior in the lower atmosphere and ionospheric absorption.

REFERENCES

- Allington, K., B. W. Boville, and F. K. Hare, Midwinter ozone variations and stratospheric flow over Canada, 1958-59, *Tellus*, 12, 266-273, 1960.
- Appleton, E. V., and W. R. Piggott, Ionospheric absorption measurements during a sunspot cycle, *J. Atmospheric and Terrest. Phys.*, 5, 141-172, 1954.
- Beynon, W. J. G., and K. Davies, Simultaneous ionospheric absorption measurements at widely separated stations, *J. Atmospheric and Terrest. Phys.*, 5, 273-289, 1954.
- Beynon, W. J. G., and K. Davies, A study of vertical incidence ionospheric absorption at 2 Mc/s, *The Physics of the Ionosphere*, pp. 40-52, Physical Society Conference Report 1955.
- Byram, E. T., T. Chubb, and H. Friedman, The study of extreme ultraviolet radiation from the sun with rocket-borne photon counters, *Rocket Exploration of the Upper Atmosphere*, 1954, pp. 276-278, Pergamon Press.
- Eyfrig, R. Absorption ionosphérique extraordinaire observée en février 1952, *Ann. géophys.*, 9, 325-327, 1953.
- Gnanalingam, S., and K. Weekes, D region echoes with a radiowave of frequency 1.4 Mc/s, *The Physics of the Ionosphere*, pp. 63-70, Physical Society Conference Report 1955.
- Jones, L. M., J. W. Peterson, E. J. Schaeffer, and H. F. Schulte, Upper air density and temperature: some variations and an abrupt warming in the mesosphere, *J. Geophys. Research*, 64, 2331-2340, 1959.
- Megill, L. R., M. P. Jamnick, and J. E. Cruz, Seasonal variations in the twilight enhancement of (OI) 5577, *J. Atmospheric and Terrest. Phys.*, 18, 309-314, 1960.
- Nicolet, M., The aeronomic problem of nitrogen oxides, *J. Atmospheric and Terrest. Phys.*, 7, 152-169, 1955.
- Nicolet, M., and A. C. Aikin, The formation of the D region of the ionosphere, *J. Geophys. Research*, 65, 1469-1483, 1960.
- Stroud, W. G., W. Nordberg, W. R. Bandeen, F. L. Bartman, and P. Titus, Rocket-grenade measurements of temperatures and winds in the mesosphere over Churchill, Canada, *J. Geophys. Research*, 65, 2307-2323, 1960.
- Whitehead, J. D., The absorption of short radio waves in the ionosphere, *J. Atmospheric and Terrest. Phys.*, 10, 12-19, 1957.

(Manuscript received December 23, 1960.)

Determination of Cloud Altitude from a Satellite

RUDOLF A. HANEL

Goddard Space Flight Center
National Aeronautics and Space Administration
Washington 25, D. C.

The successful flight of Tiros I has stimulated the interest of geophysicists and meteorologists in cloud formations.

It is very difficult, if not impossible, however, to judge the type of cloud in a picture taken by a satellite. Such judgment would be greatly enhanced if the altitude of the cloud were known. A method of measuring altitudes of clouds from a satellite is proposed in this letter.

The method is based on infrared absorption by CO_2 in the air layer above clouds. The amount of air above the upper boundary of a cloud is a rapidly changing function of the altitude of the cloud. Sunlight scattered by the cloud traverses the air layer above the cloud twice before it can be detected from a satellite. The absorption by CO_2 is best measured by comparing the intensity of reflected solar radiation in a CO_2 absorption band with the intensity in an atmospheric window nearby. Since CO_2 is uniformly distributed in the atmosphere, the absorption in the CO_2 band is a measure of the amount of atmosphere above the clouds and hence of their altitude.

A similar measurement in a water vapor absorption band will yield the total amount of water vapor above the cloud, which is another parameter of meteorological interest. By the same method *Strong, Ross, and Moore* [1960] were able to detect water vapor on Venus with a balloon-borne instrument.

Simultaneously with the CO_2 and H_2O absorption, the cloud-top temperature should be registered. The same field of view should be used for all three measurements.

Preliminary investigations indicate the following wavelength bands to be best suited for the experiments:

First reference window	1.5– 1.7 μ
Water vapor absorption band	1.8– 1.9 μ
CO_2 absorption band	2 – 2.1 μ
Second reference window	2.1– 2.2 μ
Thermal emission of cloud	10 –12 μ

The simplest equipment for this experiment is a set of fixed detectors, mounted on a stabilized platform, which measure the intensities at the subsatellite point within the five spectral ranges mentioned above.

Corrections for the local solar elevation angle and wavelength dependence of back scattering of clouds must be applied. The experiment is well suited to a local noon orbit. It is restricted to the sunlit side of the earth and to moderate angles of solar incidence, perhaps to 45° or 60° .

The value of the experiment is greatly enhanced by cloud-cover pictures taken simultaneously.

REFERENCES

- Strong, John, M. D. Ross, and C. B. Moore, Some observations of the atmospheres of Venus and the earth during the Strato Lab IV balloon flight, American Geophysical Union Meeting, Washington, D. C., April, 1960.
Stroud, W. G., Initial results of the Tiros I meteorological satellite, *J. Geophys. Research*, 65, 1643–1644, 1960.

(Received December 31, 1960.)

Observed Magnetic Declinations in West Antarctica

NED A. OSTENSO AND CHARLES R. BENTLEY

*Geophysical & Polar Research Center, University of Wisconsin
Madison, Wisconsin*

During the past four austral summers the United States has maintained an active oversnow and airborne traverse program in West Antarctica and on the adjoining Ross and Filchner ice shelves. Six of these traverses included as part of their objectives periodic determinations of magnetic declination D . Declination observations were made at intervals of approximately 30 to 36 nautical miles on the oversnow traverses and at each of the seven stations occupied by the airborne traverse (Fig. 1). These measurements were made with either a Dietzgen transit or a tripod-mounted surveying compass by noting the magnetic azimuth of the center of the sun. The time of observation was obtained to the nearest 0.2 second with a chronometer whose rate was checked daily against radio station WWV. True azimuth of the sun was then computed to $\pm 0.1^\circ$ using the *Nautical Almanac* and *H. O. 214*. The values of declination obtained with the Dietzgen transit are expressed to $\pm 0.1^\circ$, with an estimated observational error of $\pm 0.2^\circ$. The surveying compass was less accurate, and the values from this instrument are expressed to $\pm 0.5^\circ$. No attempt was made to correct for temporal variations of D because, as Ostenso and Bentley [1959] have shown, most of the observation sites were too far from the magnetic observatories at Byrd and Little America stations for any correlation to be assumed. However, since all but one of the observations were made during days of character 0 or 1, these variations should introduce an error of less than 1.0° . Thus the maximum error in obtaining absolute D was taken to be 1.0° for the transit observations and 1.5° for the surveying compass observations.

The locations of all the observation sites have been plotted in Figure 1 along with isogonic lines contoured from these data at 10° intervals. For comparison, these observed isogonic lines have been superimposed on those plotted on

H. O. map 1706S (1960). Here we see a remarkably good correlation between the computed and observed isogonic lines. Such discrepancy as does exist may be in part attributed to one of two factors or a combination of both. First, the observational data were obtained over a period of three years, and secular variations occurring during this interval would result in a minor distortion of lines. Second, it is now known that the western portion of West Antarctica (i.e., roughly west of a line joining McMurdo Sound, Byrd Station, and Thurston Island) has a higher than normal magnetic susceptibility which could result in a deflection of the observed isogonic lines. Such a conclusion has been in part verified by vertical and total magnetic intensity measurements and certain geologic observations which indicate that this part of the continent is principally a volcanic province, and consequently the magnetic susceptibility is high.

Anderson [1960] analyzed four rock samples from this volcanic province and found their composition to include an average of 1.05 per cent magnetite and 4.66 per cent ilmenite. In contrast, examination of 59 specimens from the Sentinel Mountains, Mt. Johns, and Mt. Ewing, to the east of the volcanic province, shows little or no magnetite and ilmenite content. The rocks of these three areas are of a metasedimentary sequence of the green schist facies. Further, Bentley and Ostenso [in press] have graphically shown the correlation between vertical magnetic intensities and changes in sub-ice rock topography over the volcanic province whereas no such correlation occurs east of this region.

Acknowledgments. We wish to acknowledge our indebtedness to the following persons who participated in the collection of these data: W. A. Bradley, W. E. Chapman, H. A. C. Neuburg, E. C. Thiel, and F. T. Turcotte. The K and C index values from the geomagnetic observatories at Byrd and Little America stations were supplied by IGY World Data Center A, U. S. Coast and Geodetic Survey.

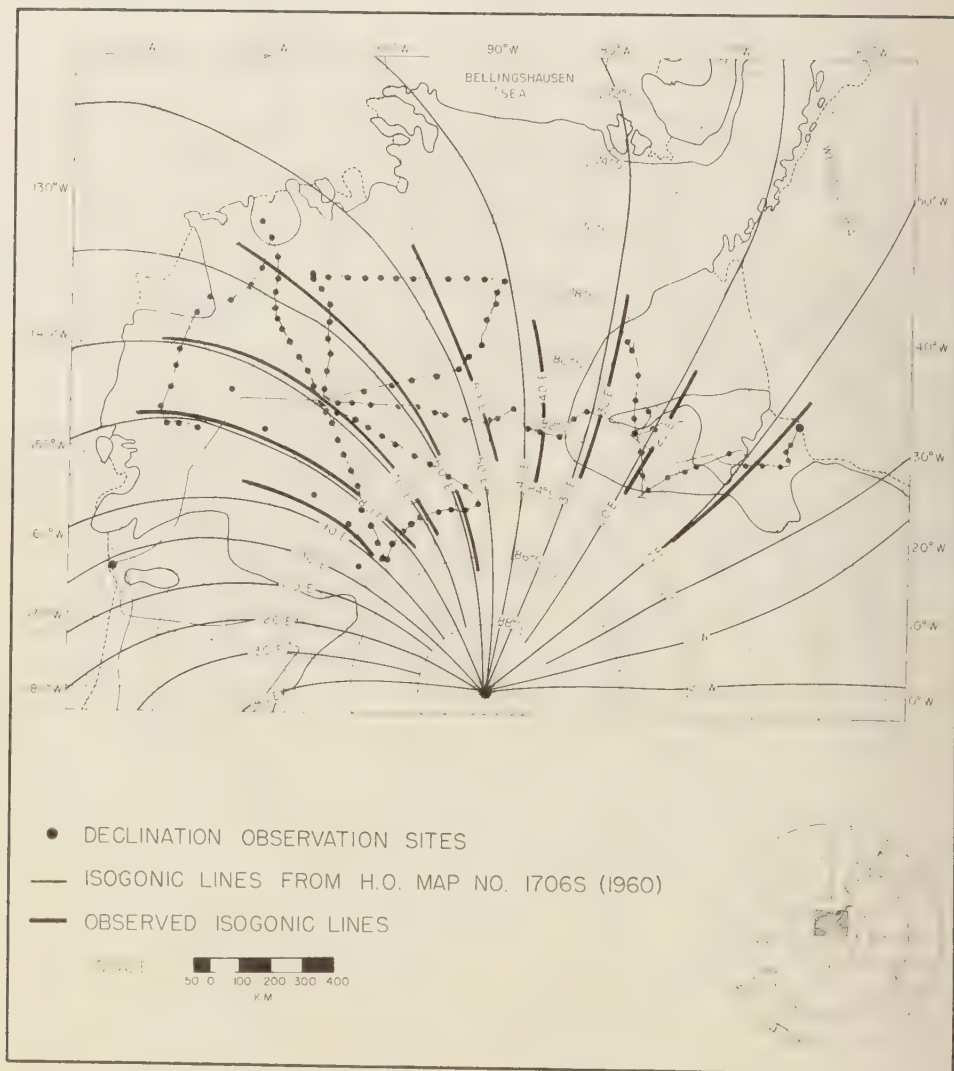


Fig. 1. Magnetic declination in West Antarctica.

REFERENCES

- Anderson, V. H., The petrography of some rocks from Marie Byrd Land, Antarctica, *The Ohio State Univ. Research Foundation Rept. 825-2—part 8*, 1960.
- Bentley, C. R., and N. A. Ostenso, Glacial and subglacial topography of West Antarctica, in press.
- Ostenso, N. A., and C. R. Bentley, Gravity and magnetic studies on the Marie Byrd Land traverses, 1957-58, *IGY Glaciological Rept. Ser. 2*, IGY World Data Center A, Glaciology, American Geographical Society, New York, July 1959.
- Thiel, E. C., and J. C. Behrendt; Gravity and magnetic measurements on the Ellsworth over-snow traverse, 1957-1958, *IGY Glaciological Rept. Ser. 2*, IGY World Data Center A, Glaciology, American Geographical Society, New York, July 1959.

(Received November 21, 1960; revised January 12, 1961.)

Photomosaic Integration of Radar Precipitation Areas

F. L. LUDWIG

*Stanford Research Institute
Menlo Park, California*

In searching for a method of graphically displaying the effects of degradation of resolution of radarscope photographs, a technique developed by *McLachlan and Chamberlain* [1959] was found to be useful for this purpose. The technique was originally proposed for use in producing mosaics from color photographs. More recently, it has come to our attention that radarscope displays, particularly when integrated photographically over a period of time, are quite useful to the hydrologist in determining rainfall amounts at points throughout a given area [Tarble, 1960; Huff, 1960]. Spatial integration would also be quite useful; the suggested mosaic technique provides a simple, rapid method of accomplishing the spatial integration from the available time-integrated transparencies.

The method involves the use of a bundle of hexagonal tubes which are reflective on their inner surfaces. The example shown here was made with aircraft structural reinforcing material manufactured by Hexcel Products, Inc. The diameter of the individual hexagonal prisms was 1/8 inch.

The principle involved is illustrated in Figure 1, which shows a thin-walled hollow cylinder with reflective walls. Translucent screens are placed at both ends. A ray of light striking the screen at end *A* is scattered in all directions, as is shown by the arrows. In some of these directions, the rays are extended to show the multiple reflections from the walls. Each of these rays generally terminates at a different point on the screen at *B*. Thus, the point illumination at the point *P* gives a uniform illumination over the entire screen at the end *B*. If colors were used, a pair of points *P* and *P'* could be illuminated by different colors, and the screen at end *B* would be uniformly illuminated with a color corresponding to the mixture of these two colors. The tube acts as an averaging device for all the illumination over the screen at *A*, producing an

even illumination at *B* corresponding to this average. This is true whether colors or gradations of gray are used. Prisms of other shapes such as the square or hexagon also display this characteristic and have the advantage of being able to cover an area completely when a number of them are fitted together.

Thus, a picture may be projected on a translucent screen at one end of a bundle of hexagonal or square prisms and be viewed on a screen at the other end in mosaic form, the brightness of each individual element corresponding to average illumination of that element of the original image.

Hydrologists are vitally interested in the total rainfall over given areas, such as river basins, during specified time intervals. Point totals of rainfall during given time intervals can be obtained from multiple-exposure pictures of a radarscope during the period [Tarble, 1960; Huff, 1960]. These have been taken, using Polaroid transparency film so that the results are immediately available at the end of the period. Presently, the spatial integration must be accomplished by isohyetal or other techniques. If the transparencies were, instead, projected through a mosaic-producing device as described, spatial integration could be accomplished by a simple summation of the rainfall equivalents of the uniform brightnesses of the individual elements of the mosaic pattern. The scale size of the elements can be adjusted by changing the projected image size.

There are two considerations involved in

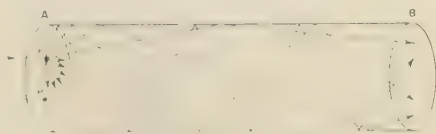


Fig. 1. Diagram showing principle involved in photomosaic construction.

choosing the scale size of the mosaic elements. First, the larger the elements, the fewer need to be summed to give the total rainfall over a given area. However, they should be small enough so that a reasonable conformation with the boundaries of the area of interest can be obtained.

The use of this technique to display visually the effects of degradation of resolution on a radarscope picture has already been mentioned and is demonstrated in Figure 2. This figure was obtained by placing a sheet of photographic paper under the 'Hexcel' material and projecting the image, with an enlarger, on the trans-



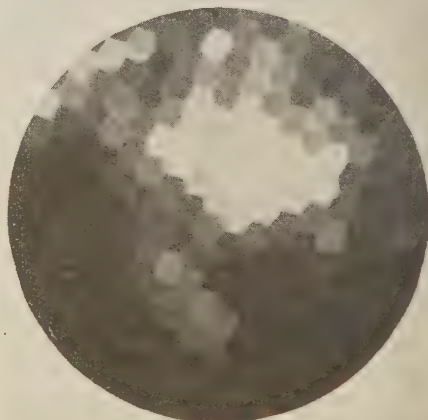
49°N 131°W
10 OCT. 1959
1940z



4 NAUTICAL MILE RESOLUTION



7 NAUTICAL MILE RESOLUTION



14 NAUTICAL MILE RESOLUTION

Fig. 2. Photomosaics of radar precipitation pattern.

ucent screen on top of the bundle of prisms. The scale of the elements was changed by use of different magnifications. These prints were then copied and reprinted, all with the same over-all dimensions. The area represented by these pictures is 200 miles in diameter.

It should be emphasized that for this type of integration the intensity of the image should be some linear function of the rainfall. This is not true for many reasons, among them being the lack of a unique relationship between radar return signal and rate of precipitation, the nonlinearities of the electronic amplification, cathode-ray-tube display, and nonlinearities in the photographic processes. However, all these difficulties are also present when other methods are used.

Obviously, use of the mosaic technique of integration would require the development of methods of calibration which could be used 'post-mosaic' as well as methods for making rapid intensity measurements and range corrections.

This technique might also be useful in the reduction of any data which can be displayed as an intensity field to grid points and digital form. However, if small intense features are of particular interest, care should be exercised in the reduction process, because such features

will appear the same, after reduction to a photomosaic, as more extensive but less intense features. Complex electronic methods are being developed for the purpose of reducing data to grid form, but the mosaic method still retains the advantages of simplicity and low cost.

The use of photomosaic techniques for integrating fields of illumination is simple and effective. Hydrological use of the technique combined with multiple image radarscope photography could speed the process of extracting, from radar data, the amount of rain which has fallen on a given area within a given period of time. This should be helpful in operational runoff forecasting; it may also be useful for limited-budget research projects.

REFERENCES

- Huff, F. A., Precipitation, *Trans. Am. Geophys. Union*, 64, 299-301, 1960.
McLachlan, Dan, Jr., and Laurie Chamberlain, SRI Memorandum on Photomosaics, (to be published elsewhere) 1959.
Tarble, R. D., The use of multiple exposure radar photographs in the Weather Bureau's hydrologic program, *Proc. Eighth Weather Radar Conf.*, San Francisco, Calif., April 11-14, 1960.

(Received November 9, 1960.)

Discussion of Paper by D. M. Hershfield and M. A. Kohler,
'An Empirical Appraisal of the Gumbel Extreme-Value Procedure'

LEO R. BEARD

*U. S. Army Engineer District
Sacramento, California*

Hershfield and Kohler [1960] merit commendation for conducting and reporting on a research project of the large scope necessary to evaluate the applicability of a statistical process. It is only through tests such as they described that precise frequency methods will be developed and agreement in methodology among the various users will eventually be reached.

There are so many tests that can be made and so many ways of interpreting test results that these matters warrant special attention. The writer hopes that the following remarks, based on studies conducted under his direction in the civil-works investigations program of the Corps of Engineers in Sacramento, will shed some additional light on this subject.

Definition of best frequency estimate. Hydrologic frequency estimating can be viewed as a process whereby one is estimating a magnitude corresponding to a specified frequency or a frequency corresponding to a specified magnitude. To define the best estimate in either case is not as simple as it might appear offhand, and the best estimate in one case does not necessarily correspond to the best estimate for the other. There are various possible definitions, including the following:

1. The magnitude that is most likely to be the correct or 'true' value.
2. The frequency that is most likely to be the 'true' frequency.
3. The magnitude or frequency that exceeds the 'true' value in half of the cases.
4. The magnitude whose successive values in a large number of cases will average the same as the average 'true' value.
5. The frequency whose successive value in a large number of cases will average the same as the average 'true' frequency.

The principal test employed by the authors (averaging observed frequencies) is a validity test only for estimates based on definition 5.

However, it is not apparent that the frequency estimates were made in accordance with that definition (it would be well if the authors would specify the fitting procedure used in their study), and it is believed that this might explain why the test results differ consistently for different sample sizes.

The writer has discussed a procedure for estimating frequencies or probabilities under definition 5 and the desirability of using that definition [Beard, 1960]. He demonstrated that the frequency or probability estimate must be made somewhat differently for different sample sizes (as suggested by Helmert's 3-sigma fallacy discussed by Gumbel in the authors' sixth reference). A frequency or probability estimate based on that adopted definition is called the expected frequency or expected probability, because it conforms to the concept of mathematical expectation.

Since the authors' test might not conform to the theory on which the frequency estimates are based, and because the expected frequency estimate would then be much greater than the authors' estimate in the range of rare frequencies, there is some question whether the extreme-value function does not in fact specify rainfall frequencies that are larger than actually occur. To check this, the writer has constructed 64 random samples, each containing 25 events drawn by the use of random numbers from the Fisher-Tippett extreme-value function. Frequency estimates were based in turn on 10 of these in each sample and on the entire 25, and the computed frequencies were compared with observed frequencies in an independent 25-event sample. The expected and observed frequencies were then added for all 64 samples. It is believed that this procedure is identical to that used by the authors for the construction of their Table 3, except that the number of samples used herein is exactly half of the 128 used by

TABLE 1. Comparison of Expected and Observed Frequencies

Criteria for Infinite Sample Size		Expected-Probability Criteria					
		10-Event Samples			25-Event Samples		
		Exp. Freq.	Observed		Exp. Freq.	Observed	
N-Year Event	Exp. Freq.		Rain*	Random		Rain*	Random
2	1600	1600	1597	1594	1600	1570	1554
5	640	714	581	756	685	570	664
10	320	406	323	434	352	288	378
25	128	208	141	238	160	113	170
50	64	132	79	152	90	61	88
100	32	86	43	92	51	37	42
200	16	58	21	64	29	18	26
500	6.4	36	12	38	15	4 _n	12
1000	3.2	26	9	32	9	3	8
5000	0.6	13	4	16	3	1	4
10,000	0.3	10	2	8	2	1	2

* Obtained from *Hershfield and Kohler* [1960, Table 3].

the authors. The expected and observed frequencies were therefore doubled in order to make a direct comparison with the authors' results. The comparison is given in Table 1.

The extreme-value function was fitted to the random samples by use of moments (mean and standard deviation) of the values. The expected frequencies were estimated from comparison criteria derived for the normal distribution and applied to the extreme-value function. These are therefore approximate, but are apparently reasonably close.

Since the observed rain events are far fewer than expected theoretically or than experienced by events known to conform to the extreme-value function, it would appear that they cannot reasonably be assumed to have derived from populations described by the extreme-value

function. Perhaps the authors' method of fitting can account for the apparent divergence.

It is interesting to note that the ratios of rain frequencies observed by the authors for 10-event and 25-event samples approximate the ratios of the theoretical expected frequencies and are therefore highly suggestive of the validity of the expected-probability concept.

REFERENCES

- Hershfield and Kohler, An empirical evaluation of the Gumbel extreme-value procedure, *J. Geophys. Research*, 65, 1737-1746, 1960.
 Beard, L. R., Probability estimates based on small normal-distribution samples, *J. Geophys. Research*, 65, 2143-2148, 1960.

(Received September 16, 1960; revised November 17, 1960.)

Authors' Reply to Discussions by I. I. Gringorten,
M. A. Benson, and L. R. Beard of the Paper,
'An Empirical Appraisal of the Gumbel Extreme-Value Procedure'

D. M. HERSHFIELD AND M. A. KOHLER

*Hydrologic Services Division, U. S. Weather Bureau
Washington, D. C.*

Gringorten [1960] made the interesting point that an equation of estimate should perform better upon the data sample from which it was derived than it would upon an independent sample. Consideration should be given both to our method of analysis and to the type of data analyzed to show why this was not the case. It is not unusual in a series of extreme rainfalls to have one or more anomalous events. An analysis by the Gumbel procedure which includes this event (or events) attaches a larger probability to it than an analysis which does not include the anomalous event. The 12.12-inch 24-hour rainfall at Hartford, Conn., in 1955 is a case in point. This rainfall, the largest ever observed at Hartford, was nearly twice as large as the previously observed maximum. A frequency analysis which does not include this event attaches a 10,000-

year return period to it. If the event is included in the analysis, it becomes a 1000-year value. By separating each station's data into two parts, we underestimated in some instances and overestimated in others, but we felt that with a large sample the opposing errors would cancel each other.

Gringorten suggested that the simplified equation

$$x = \bar{x} + s_z(y - 0.5772)/1.28255$$

which does not depend on the number of years of record, might provide better estimates than Gumbel's equation of estimate. Our experience with rainfall extremes indicates that some adjustment must be made for length of record because both the sample mean and the standard deviation generally increase with increasing length of record. For example, an examination of the standard deviations from 50 stations computed for both a 10- and a 50-year record shows that the average difference is approximately 20 per

cent and that 34 of the larger standard deviations come from the longer records.

We agree with *Benson* [1960] that by χ^2 standards the results of testing with the independent data do not show a good fit to the Fisher-Tippett type I distribution when the Gumbel procedure is used. However, we cannot agree that the New England data show consistently concave upward curves on extreme probability paper. Visual inspection of the rainfall extremes for three durations for 11 Weather Bureau stations in New England (Table 1) reveals the behavior of the data on extreme probability paper.

There can be no quarrel with the statement from *Beard's* [1961] discussion that there are many tests that can be made and also many

TABLE 1. Behavior of Maximum Rainfalls on Extreme-Probability Paper

Station	Period of Rec.	Duration		
		10-Min	60-Min	24-Hours
Hartford, Conn.	1905-57	L*	L	CU†
New Haven, Conn.	1905-57	CU	L	CU
Eastport, Maine	1903-52	L	L	L
Portland, Maine	1903-21	L	L	L
Boston, Mass.	1897-57	L	L	CU
Nantucket, Mass.	1903-57	CD‡	L	CD
Concord, N. H.	1905-57	L	L	L
Block Is., R. I.	1903-48	L	L	CD
Providence, R. I.	1905-57	L	L	L
Burlington, Vt.	1906-57	L	L	L
Northfield, Vt.	1903-43	L	L	L

* L—Linear.

† CU—Concave upward.

‡ CD—Concave downward.

ways to evaluate the results of these tests. In appraising a procedure, one can measure its value in terms of what was attempted. We attempted to evaluate the Gumbel procedure by viewing the data from a large number of stations in the aggregate, thereby substituting space for time. The analysis was made within the framework of certain assumptions which were made explicit. The fact that Beard's results do not agree with ours can be ascribed to the difference in the data analyzed. Beard used hypothetical data manufactured from the Fisher-Tippett type I distribution. We used the actual rainfall observations and attempted to determine whether the combination of Gumbel's curve-fitting procedure and the Fisher-Tippett type I distribution, which is

strongly supported by theory, would be acceptable for practical purposes.

REFERENCES

- Beard, L. R., Discussion of paper by D. M. Hershfield and M. A. Kohler, 'An empirical appraisal of the Gumbel extreme-value procedure,' *J. Geophys. Research*, 66, 1306-1307, 1961.
- Benson, M. A., Discussion of paper by D. M. Hershfield and M. A. Kohler, 'An empirical appraisal of the Gumbel extreme-value procedure,' *J. Geophys. Research*, 65, 4219, 1960.
- Gringorten, I. I., Discussion of paper by D. M. Hershfield and M. A. Kohler, 'An empirical appraisal of the Gumbel extreme-value procedure,' *J. Geophys. Research*, 65, 4220, 1960.

(Received January 25, 1961.)

Discussion of Paper by P. P. Rowe, 'An Equation for Estimating Transmissibility and Coefficient of Storage from River-Level Fluctuations'

MAHDI S. HANTUSH¹

*New Mexico Institute of Mining and Technology
Socorro, New Mexico*

In this paper Rowe [1960], dealing with an important problem in ground-water hydrology, presents solutions for the fluctuation of ground-water levels in response to water-level fluctuations in streams that cut through the water-bearing materials. He treats three cases in which the stream level varies with time t in accordance with the relations ct , $c_1\sqrt{t}$, and $c_2t(t_1 - t)$, respectively. The solutions are presented in the form of equations 11, 17, and 20, of which (11) is evaluated in the referenced paper. It appears, however, that C_2 of (9), which Rowe has evaluated as zero, should have the value $2ctU^2$. In this case (11) should be

$$h = ct[1 + 2U^2 - (4/\sqrt{\pi}) \operatorname{erf}(U)] \quad (11a)$$

The same correction would apply to (13), (14), and (15).

The corrected value of C_2 is obtained as follows:

The series representation of $\operatorname{erf}(U)$ can be put in a closed form as

$$\begin{aligned} \frac{4}{\sqrt{\pi}} \operatorname{erf}(U) &= -\frac{4}{\sqrt{\pi}} \sum_{n=0}^{\infty} \frac{(-1)^n U^{2n+1}}{n!(2n-1)(2n+1)} \\ &= \frac{2}{\sqrt{\pi}} \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \left(\frac{1}{2n+1} \right. \\ &\quad \left. + 1 - \frac{2n}{2n-1} \right) U^{2n+1} \\ &= \left(\frac{2}{\sqrt{\pi}} \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \frac{U^{2n+1}}{(2n+1)} \right) \\ &\quad + \frac{2U}{\sqrt{\pi}} \left(\sum_{n=0}^{\infty} \frac{(-1)^n}{n!} U^{2n} \right) \end{aligned}$$

$$\begin{aligned} &+ 2U^2 \left(\frac{2}{\sqrt{\pi}} \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{(n-1)!(2n-1)} \frac{U^{2n-1}}{(2n-1)} \right) \\ &= \operatorname{erf}(U) + \frac{2U}{\sqrt{\pi}} \exp(-U^2) \\ &\quad + 2U^2 \operatorname{erf}(U) \\ &= (1 + 2U^2) \operatorname{erf}(U) \\ &\quad + \frac{2U}{\sqrt{\pi}} \exp(-U^2) \end{aligned}$$

where $\operatorname{erf}(U)$ is the error function.

With the above relation, equation 9 of the original paper becomes

$$\begin{aligned} h = c \left\{ t - \left(t + \frac{x^2}{2\nu} \right) \operatorname{erf}(U) \right. \\ \left. - \frac{2}{\sqrt{\pi}} \frac{x\sqrt{t}}{\sqrt{4\nu}} \exp(-U^2) \right\} + C_2 \end{aligned}$$

in which $\nu = T/S$.

For $t = 0$: $h(x, t) = 0$, $U = \infty$, $\operatorname{erf}(\infty) = 1$ and $\exp(-\infty) = 0$. It follows, therefore, that $C_2 = cx^2/2\nu = 2ctU^2$.

The solution given by (11a) involves the function erf , which is tabulated briefly in Table 1 of the original paper. The values of the parameter U encountered in practice vary from 0 to about 3. This would require the extension of Rowe's table. It may be more convenient to put the solutions for the three cases under discussion in terms of function that have already been tabulated for a wide range of the parameter U . These are obtained as follows:

It is shown [Carslaw and Jaeger, 1947, p. 382 items 10 and 11] that the inverse Laplace transform of the function $(1/p^2) \exp(-x\sqrt{p/\nu})$ is given by

¹ On leave from University of Baghdad, Iraq.

$$\begin{aligned}
 & L^{-1} \left[\frac{1}{p^2} \exp \left(-x \sqrt{\frac{p}{\nu}} \right) \right] \\
 &= \left\{ \left(1 + \frac{2x^2}{4\nu t} \right) \operatorname{erfc} \left(\frac{x}{\sqrt{4\nu t}} \right) \right. \\
 &\quad \left. - \frac{2}{\sqrt{\pi}} \frac{x}{\sqrt{4\nu t}} \exp \left(-\frac{x^2}{4\nu t} \right) \right\} \quad (21) \\
 &= 4t \, i^2 \operatorname{erfc} \left(\frac{x}{\sqrt{4\nu t}} \right) \quad (22)
 \end{aligned}$$

In general

$$\begin{aligned}
 & L^{-1} \left[\frac{1}{p^{i+n/2}} \exp \left(-x \sqrt{\frac{p}{\nu}} \right) \right] \\
 &= (4t)^{n/2} i^n \operatorname{erfc} (x/\sqrt{4\nu t}) \quad (23)
 \end{aligned}$$

in which $n = 0, 1, 2, 3, \dots$. Also $i^n \operatorname{erfc}$ is the n th repeated integral of the complementary error function and is defined by

$$i^n \operatorname{erfc} (y) = \int_y^\infty i^{n-1} \operatorname{erfc} (\beta) d\beta$$

with $i^0 \operatorname{erfc}(y) = \operatorname{erfc}(y) = 1 - \operatorname{erf}(y)$.

By using (3) of the original paper, and with the aid of (21), (22), and (23), the solution to the three cases, as well as to the general case in which $h(0, t) = f(t)$, will, in the notation of Rowe, be found readily as

1. For $h(0, t) = ct$, which transforms into $\bar{h}(0, p) = c/p^2$:

$$\begin{aligned}
 h(x, t) &= ct[4i^2 \operatorname{erfc} (U)] \\
 &= ct \{ (1 + 2U^2) \operatorname{erfc} (U) \\
 &\quad - (2/\sqrt{\pi}) U \exp (-U^2) \} \quad (24)
 \end{aligned}$$

2. For $h(0, t) = c_1 \sqrt{t}$, which transforms into $\bar{h}(0, p) = (c_1 \sqrt{\pi})/(2p^{3/2})$:

$$\begin{aligned}
 h(x, t) &= (c_1 \sqrt{\pi t}/2) [2i \operatorname{erfc} (U)] \\
 &= c_1 \sqrt{t} \{ \exp (-U^2) \\
 &\quad - \sqrt{\pi} U \operatorname{erfc} (U) \} \quad (17)
 \end{aligned}$$

3. For $h(0, t) = c_2 t(t_1 - t)$, which transforms into $\bar{h}(0, p) = c_2 t_1/p^2 - 2c_2/p^3$:

$$\begin{aligned}
 h(x, t) &= c_2 t \{ t_1 [4i^2 \operatorname{erfc} (U)] \\
 &\quad - 4t [8i^4 \operatorname{erfc} (U)] \} \quad (26)
 \end{aligned}$$

4. For $h(0, t) = f(t)$ which transforms into $\bar{h}(0, p) = \bar{f}(p)$:

As the inverse Laplace transform of the function $\exp (-x\sqrt{p/\nu})$ is given by $(x/2\sqrt{\pi \nu t^3}) \exp (-x^2/4\nu t)$, the solution is

$$\begin{aligned}
 h(x, t) &= \frac{x}{2\sqrt{\pi \nu}} \int_0^t f(\tau) \exp \left(-\frac{x^2}{4\nu(t-\tau)} \right) \\
 &\quad \cdot \frac{d\tau}{(t-\tau)^{3/2}}
 \end{aligned}$$

$$= (2/\sqrt{\pi}) \int_U^\infty f(t - x^2/4\nu\beta^2) e^{-\beta^2} d\beta \quad (27)$$

of which (24), (25), and (26) are special cases.

Tables for the function $i^n \operatorname{erfc}$ are available in the literature [Carslaw and Jaeger, 1947, p. 373, or 1959, p. 485; also Kaye, 1955, p. 119].

The second term in the bracketed series in the equation following (9) should read $\frac{1}{3}U^2$, and T in (13) should be changed to t ; these are obviously typographical errors.

REFERENCES

- Carslaw, H. S., and J. C. Jaeger, *Conduction of Heat in Solids*, Oxford Univ. Press, London and New York, first edition, 1946, second edition, 1959.
- Kaye, J., A table of the first eleven repeated integrals of the error function, *J. Math. Phys.*, 34, 119-125, 1955.
- Rowe, P. P., An equation for estimating transmissibility and coefficient of storage from river-level fluctuations, *J. Geophys. Research*, 65, 3419-3424, 1960.

(Received November 21, 1960.)

API Nomogram

WAYNE J. KAMMERER¹*California Department of Water Resources
Sacramento, California*

The Antecedent Precipitation Index nomogram was designed as a computational aid for the hydrologist who uses 0.9 Antecedent Index [Kohler and Linsley, 1951; Linsley, Kohler, and Paulus, 1958]. Especially useful during the development of flood forecast procedures, it can be used for API computations during rainfall-runoff forecasting. It is general in that it can be used with precipitation data from any station, and its upper limit is believed to be high enough for nearly any area in the United States.

By use of the nomogram, the 0.9 factor and the addition of the daily rainfall are included in a single step. For example, to compute today's API:

If there has been no rainfall in the past 24 hours, place a straightedge on the last computed API on the left-hand scale and follow diagonally to the right through the point which lies on the line of zero additional rain as well as on the line which represents the correct number of rainless days. (Thus, if the API on the 5th is 10.0, if there has been no rain on the 6th or 7th, and if

we wish to compute the API on the 8th, lay the straightedge on 10.0 on the left-hand scale, follow diagonally to the right through the intersection of the line of zero additional rain with '2 rainless days between, etc.,' and read 7.3 as the new API on the 8th.)

To compute the API on the 9th, when there is an additional 1.5 inches of rain on the 9th set the straightedge at 7.3 on the left-hand scale, follow diagonally to the right through the intersection of the line of 1.5 inches additional rain with 'rainless days between, etc.,' and read 8.3 as the API on the 9th.

(Full size copies of the nomogram may be secured from Delbert D. McNealy, Department of Water Resources, P. O. Box 388, Sacramento, California.)

REFERENCES

- Kohler, M. A., and R. K. Linsley, Predicting the runoff from storm rainfall, *U. S. Weather Bureau Research Paper 34*, 1951.
Linsley, R. K., M. A. Kohler, and J. L. H. Paulus, *Hydrology for Engineers*, McGraw-Hill Book Co. New York, 171-172, 1958.

(Received December 7, 1960.)

¹ Present address: Portland Cement Association, Champaign, Illinois.

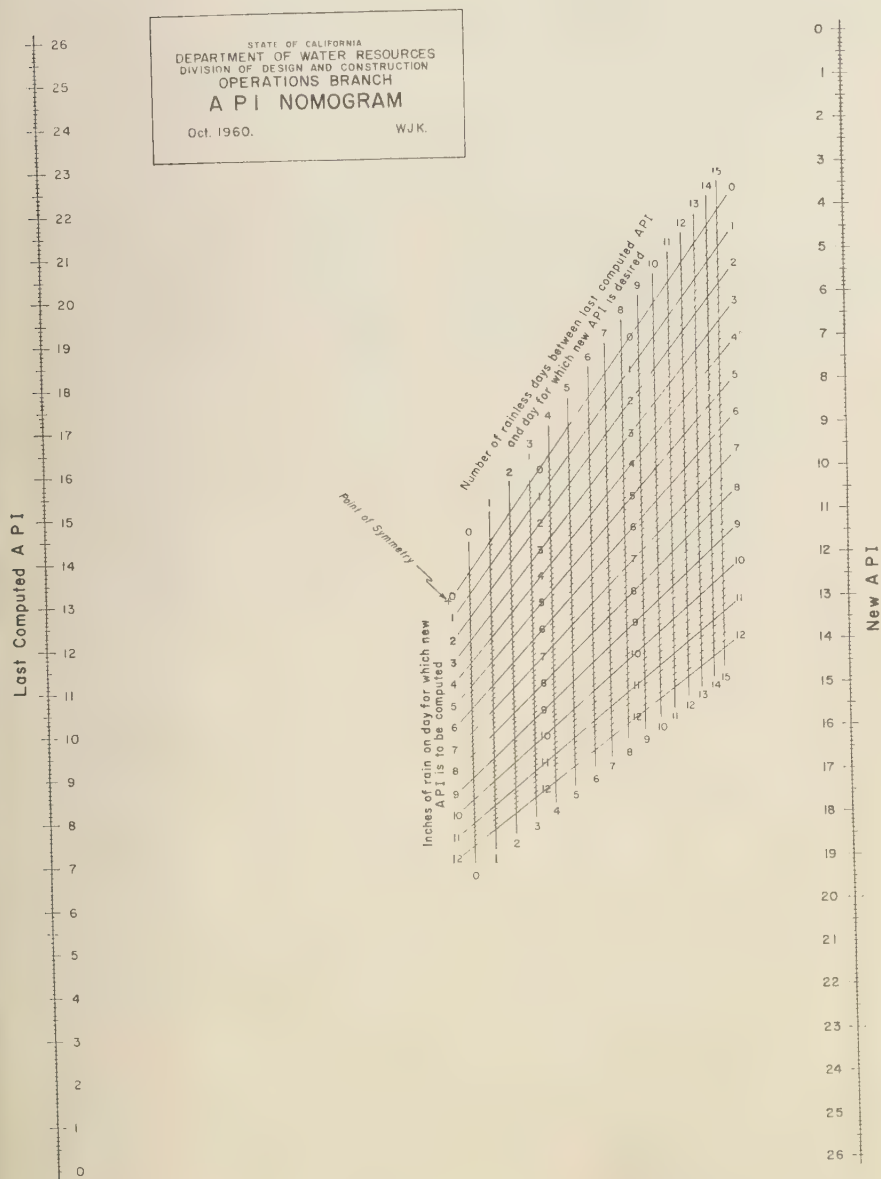


Fig 1. API nomogram.

Seasonal Variations in Atmospheric Carbon Dioxide Concentration

EUGENE M. WILKINS

*Temco Electronics and Missiles Company
Dallas, Texas*

The importance of carbon dioxide to the heat budget of the atmosphere is now well known. The present indication is that the carbon dioxide is increasing by the amount added by fossil fuel combustion (no removal process), and this situation can be expected to result in a gradual change of the world's climate [Plass, 1956]. If 6×10^9 tons of carbon dioxide are assumed to be added annually, the annual increase in concentration should be about 1.4 ppm [Revelle and Suess, 1957], which is very near the values currently observed [Keeling, 1960a], namely 1.3 ppm at the south pole and 0.5 to 1.2 ppm over the northern Pacific Ocean.

Because of CO_2 absorption of infrared radiation the secular increase in concentration should result in a world-wide warming at the rate of about 1° to 2°F per century [Kaplan, 1960]. One of the most important consequences of this warming is that there might be a gradual rise in ocean level due to melting of polar ice caps which could eventually cause inundation of several populous coastal regions such as Long Island, Florida, and Holland.

The first systematic CO_2 research program producing meaningful results has been under the direction of C. D. Keeling at Scripps Institution of Oceanography since the beginning of the IGY. Some excellent presentations of the data have been given by Keeling and Rakestraw [1960] and by Keeling [1960a, b]. The care exercised in the gathering of data samples and the precise techniques for concentration analysis have provided data of excellent quality. Thus, as the program continues, we may expect to see several more gaps filled in our knowledge about this important aspect of atmospheric physics and world climate.

The present interest is in the interpretation of the seasonal variation of CO_2 concentration and the possible implications on the amount of mixing between the hemispheres.

Keeling has suggested that the rather large seasonal range of CO_2 concentration in the northern hemisphere may result from the activity of plant life on the land areas. Maximum concentrations are seen to occur in the spring, at the outset of the growing season, and minimum concentrations occur in the fall, near the end of the growing season. A further indication that plant life is largely responsible for this variation is seen in the relative abundance of the isotope C^{13} and C^{12} near plants, since the ratio of these isotopes has a similar seasonal trend.

We note, however, that the seasonal variation in the combustion of fossil fuels for heating would be expected to follow a similar trend, namely a steady build-up during the colder months and a decline during the summer. If a significant secular increase, which currently is apparent in the data, is attributed to this activity of mankind, it follows that the same activity can be expected to appear in the seasonal variation also, though doubtless to a lesser degree.

Keeling has also suggested that the absence of a seasonal trend in the southern hemisphere may be due to the relative scarcity of plant life there in temperate and polar latitudes. However, in this suggestion he does not consider whether a significant exchange of air occurs between the two hemispheres. When this is considered, it is possible to offer an alternative explanation for the observed seasonal variations in CO_2 concentration.

If the seasonal effects were insignificant in the southern hemisphere, exchange between the hemispheres would result in a regime of small amplitude in the southern hemisphere, but almost in phase with that of the northern hemisphere. The exchange is strongly indicated by the observation of a rather large secular increase of CO_2 at Antarctica.

If, on the other hand, there is an appreciable effect in the southern hemisphere attributable

to the combined activities of man and plants, then in the absence of interhemispheric exchange there should be a noticeable seasonal variation, of smaller amplitude and of *opposite phase* with that of the northern hemisphere. According to the findings in the northern hemisphere, this regime should be just as pronounced in the Antarctic, where the measurements were made, as in the temperate latitudes.

The more reasonable conclusion appears to be that the regimes in the two hemispheres, being out of phase, tend to cancel each other. The larger amplitude in the northern hemisphere prevails, and the smaller one in the southern hemisphere is canceled out in the mixing. Thus, except for the exchange, the seasonal regime in the northern hemisphere would have an even larger amplitude than is presently observed.

The human and plant life activities that would be expected to contribute most to the seasonal range of CO_2 concentration are found between latitudes 20° and 60° . The southern hemisphere has about one-fourth as much land area between these latitudes, indicating that the seasonal effect there should be about one-fourth as great as that of the northern hemisphere. Assuming that interhemispheric mixing occurs, and using the observed seasonal ranges of about 6.5 ppm for the northern hemisphere and 0 ppm for the southern hemisphere, we can compute that in the absence of mixing the seasonal range should be about 8.6 ppm in the northern hemisphere and about 2.15 ppm in the southern hemisphere. These figures lead to a calculation which shows that about one-fifth of the air in one hemisphere effectively mixes with that of the other hemisphere every half year.

According to a recent communication from Dr. Keeling, carbon dioxide measurements are in progress at mid-latitudes in the southern hemisphere. If the new data do not show a significant seasonal variation, we may regard as reasonable the above estimate of the interhemispheric exchange rate. If a significant seasonal variation occurs, it will prove difficult to reconcile the variation with the absence of one in the Antarctic, since lateral mixing should mask such an effect.

Acknowledgment. I wish to express my thanks to Dr. Charles D. Keeling for much enlightenment in the field of atmospheric carbon dioxide research, in assistance with this work as well as in our pleasurable association during the IGY.

REFERENCES

- Kaplan, L. D., The influence of carbon dioxide variations on the atmospheric heat balance, *Tellus*, 12, 204-208, 1960.
Keeling, C. D., The concentration and isotopic abundances of carbon dioxide in the atmosphere, *Tellus*, 12, 199-203, 1960a.
Keeling, C. D., Measurements of carbon dioxide in the atmosphere (IGY Bulletin), *Trans. Am. Geophys. Union*, 41, 512-515, 1960b.
Keeling, C. D., and N. W. Rakestraw, The concentration of carbon dioxide in the atmosphere (Abstract), *J. Geophys. Research*, 65, 2502, 1960.
Plass, G. N., The carbon dioxide theory of climatic change, *Tellus*, 8, 140-154, 1956.
Revelle, R., and H. E. Suess, Carbon dioxide exchange between atmosphere and ocean and the question of an increase of atmospheric carbon dioxide during the past decades, *Tellus*, 9, 18-27, 1957.

(Received December 1, 1960; revised January 10, 1961.)

Addendum: The Density and Mass Distribution of Meteoritic Bodies in the Neighborhood of the Earth's Orbit¹

HARRISON BROWN

*California Institute of Technology
Pasadena, California*

In an earlier paper [Brown, 1960], the author estimated the frequency of meteorite impact upon the earth and moon on the basis of the numbers of observed falls over a period of a century in Japan, India, and Western Europe. All of these areas have had high rural population densities during the entire period. It was recognized, however, that the estimated fall density of 0.32 falls/year/10°/km² was probably low.

Recently, in connection with the compilation of maps showing the locations of observed meteorite falls, it has been noticed that approximately $\frac{1}{3}$ of the Indian falls have been recovered in the relatively small area of 330,000 km² embracing most of the United Provinces, the eastern parts of the Punjab and Rajputana, and the western part of Bihar Province. This region is characterized by the large river network that includes the Ganges, the Jumna, the Gogra, the Chauka, the Son, and the Gandak. The region is extremely fertile and level, with the result that it has one of the highest rural population densities in the world—more than 500 persons per square mile.

During the half-century 1860–1909 twenty

meteorites were observed to fall in this region. In the following half-century (1910–1959) an additional 13 were observed, bringing the total for one century to 33. The annual rate of fall per unit area in this region corresponds, then, to 1.0 falls/year/10°/km². This is higher than the earlier estimate by a factor of 3.

It is noted that there are similar concentrations of falls in other parts of the world. Correcting the Japanese data for the fact that most falls there have been observed in the southern regions the Japanese rate of fall per unit area appears to be identical with that in the United Provinces of India within statistical error. A similar high rate has prevailed in the Po Valley in northern Italy where 7 meteorites have been observed to fall in the past 150 years in an area of only 43,000 square kilometers.

The observed rates of fall in the regions of unusually high fall density in India and Japan are summarized in Table 1.

Using the term 'fall' to denote a meteorite that passes through the atmosphere and, after landing, is large enough to be found and picked up, the total rate of fall upon the earth appears to be about 560 meteorites per year. In view of the higher value of the estimated fall density upon the earth, all of the computed impact frequencies of meteorites upon the earth and

¹ Contribution No. 1021, Division of Geological Sciences, California Institute of Technology, Pasadena, California.

TABLE 1. Meteorite Falls in Selected Regions, 1860–1959

Region	Area 10°/km ²	Number		Falls/year/10°/km ²	
		1860–1909	1910–1959	1860–1909	1910–1959
North-central India	330	20	13	1.2	0.8
Southern Japan	190	11	10	1.2	1.1
Total	520	33	23	1.3	0.9
Average for century				1.1	

noon, given in Table 3 [Brown, p. 1682, 1960]
should be multiplied by a factor of 3.4.

Acknowledgments. This work was supported by the National Aeronautics and Space Administration under Grant NsG-56-60. I am indebted to Mr. Gregory Smith and Mrs. Eleanor Helin who helped with the plotting of the data.

REFERENCES

- Brown, Harrison, The density and mass distribution of meteoritic bodies in the neighborhood of the earth's orbit, *J. Geophys. Research*, 65, 1679-1683, 1960.

(Received January 19, 1961.)

Note on 'The Economical Net Radiometer'

C. B. TANNER

*Department of Soils, University of Wisconsin
Madison, Wisconsin*

J. A. BUSINGER

*Department of Meteorology, University of Washington
Seattle, Washington*

AND

P. M. KUHN

U.S. Weather Bureau, Madison, Wisconsin

After the submission of our manuscript, 'The Economical Net Radiometer,' which appeared in the November 1960 issue of this Journal, a similar derivation for the *black surfaced* radiometer by D. W. Scholte Ubing came to our attention.

Scholte Ubing's work referred to is 'Studies on solar and net radiation and on evapotranspiration of grass,' *Mededel Landbouwhogeschool Wageningen*, 59 (10), 1-93, 1959 (in Dutch).

(Received November 21, 1960.)

Comment on Paper by D. G. Singleton, 'The Geomorphology of Spread F' '

JOHN P. MILLER

Harvard University, Cambridge, Massachusetts

Singleton's use of the word 'geomorphology' seems to extend this well-established, rapidly growing branch of geology and geography into a domain not previously claimed by even its most imperialistic practitioners.

Reading the title, I guessed immediately that the paper must be concerned with the landforms of a huge Texas ranch, but on turning to the article and noting the author's Australian address, I concluded that the locale must be somewhere in the 'Out Back.' Needless to say, these logical deductions were demolished by the

first few lines of an abstract dealing with the distribution of isopleths of certain ionosonde data.

As presently defined, geomorphology is devoted to study of the earth's surface configuration. In my opinion, an earth-bound view of this subject is not outmoded in the Space Age.

REFERENCE

Singleton, D. G., The geomorphology of spread F , *J. Geophys. Research*, 65, 3615-3624, 1960.

(Received January 6, 1961.)

Corrigenda

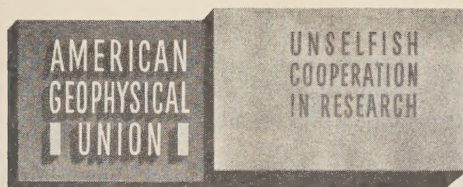
Mr. R. A. Duncan has called attention to an error in his note, 'The Cause of Magnetic Storms and Bays,' published in vol. 65, no. 11, of this *Journal*. Page 3589 gives the wrong direction to the Hall current resulting from electric fields in the ionosphere, thus nullifying the argument that follows. Unfortunately this error was noticed too late to stop publication.

Attention has been called to an error in Figure 1 of 'Potential Evapotranspiration Estimates by the Approximate Energy Balance Method of Penman,' by C. B. Tanner and W. L. Pelton,

p. 3408 of the October 1960 issue of the *Journal*. The inscriptions on the two curves should be interchanged so that the dashed line is labeled 'vapor pressure' and the solid line ' $\Delta e/\Delta T$.'

• Professor Mahdi S. Hantush has called attention to an error in his paper, 'Modification of the Theory of Leaky Aquifers,' published in the November 1959 issue of the *Journal*. On page 3718, first column, the line following equation 18 should be

$$\text{for } u \leq \text{than both } 10^{-5}/\beta^2 \text{ and } 10^{-4}\beta^2.$$



CORPORATION AND SUPPORTING MEMBERSHIPS

The American Geophysical Union is a non-profit scientific organization established by the National Research Council. Its Council is the United States National Committee of the International Union of Geodesy and Geophysics, through official adherence by the United States through the National Academy of Sciences-National Research Council.

Extracts from the Statutes:

Article 3. Membership—The membership of the American Geophysical Union shall be as follows:

- (e) *Corporation Members*—Corporations and other organizations interested in geophysics elected by the Executive Committee of the Union. The designated representative of each such organization shall enjoy the privileges of a Member.
- (g) *Supporting Members*—Corporations, other organizations, and individuals interested in geophysics and desirous of supporting the Union may become members under the following classifications upon election by the Executive Committee

(Continued on next page)

Cut along this line

American Geophysical Union

PROPOSAL FOR _____ MEMBERSHIP

To the Executive Committee, American Geophysical Union
1515 Massachusetts Avenue, N. W., Washington 5, D. C.

Gentlemen:

As an indication of our interest in the aims and activities of the American Geophysical Union, and to assist in maintaining and extending its program of publication and other work in the development of the geophysical sciences, the undersigned applies for _____ Membership in the AGU and, until further notice, agrees to pay annual dues at the rate established for this classification of membership, in accordance with the information set forth above and on the following page.

Corporation or Organization _____

By _____ Title _____

(Signature)

(over)

(Continued from previous page)

of the Union . . . : Contributing Members, . . . Sustaining Members, . . . Benefactors. . . . Each Supporting Member and the designated representative . . . shall enjoy the privileges of a Member.

Corporation Members shall pay dues of not less than \$100 for each calendar year. Dues for Supporting Members shall be as follows:

Contributing Members	\$500
Sustaining Members	\$1000
Benefactors	\$5000

Lists of Corporation Members and Supporting Members will be published in each issue of the *Transactions*, and will be included in the Membership Directory as distinct units.

By-Laws provide that one copy of each issue of the *Transactions*, *Journal of Geophysical Research*, any published *List of Members and Officers*, and any other publication which may be approved for free distribution to the membership shall be sent to each Corporation and Supporting Member. Each organization in good standing may purchase any available publication of the Union at the established member discount.

AMERICAN GEOPHYSICAL UNION

1515 Massachusetts Ave., N.W.
Washington 5, D. C.

Cut along this line

Address _____

City _____ State _____

General fields of activity _____

The following person is designated as our representative in this membership _____

_____ Title _____

Number of units of membership desired (this will be taken as one unless otherwise indicated) _____

Place _____

Date _____

Contents

(Continued from back cover)

	PAGE
Edibility of Some California Wildland Soils Related to Their Metallic Cation Exchange Capacity <i>James R. Wallis and Lee J. Stevan</i>	1225
Heat Flow from a Differentiated Earth..... <i>Sydney P. Clark, Jr.</i>	1231
Geomagnetic Study of the Sudbury Basin..... <i>Peter J. Hood</i>	1235
Earth Currents of Deep Internal Origin..... <i>P. H. Roberts and F. J. Lowes</i>	1243
Cosmic Profiles in Northwestern Utah: Pilot Range and Grouse Creek Range Area <i>Joseph W. Berg, Jr., Kenneth L. Cook, Harry D. Narans, Jr., and Richard J. Leamer</i>	1255
Gravity Anomalies and Crustal Section across the Tonga Trench <i>Manik Talwani, J. Lamar Worzel, and Maurice Ewing</i>	1265
Geomagnetic and Solar Data..... <i>J. Virginia Lincoln</i>	1279
Letters to the Editor:	
Short-Term Variations in Meson and Nucleon Component of Cosmic Rays <i>V. L. Patel and K. Maeda</i>	1286
A Note on Solar Flare Cosmic Rays..... <i>K. Maeda and V. L. Patel</i>	1288
Solar-Stream Distortion of the Geomagnetic Field and Polar Electrojets..... <i>J. W. Kern</i>	1290
Some Results of Direct Probing in the Ionosphere <i>W. Pfister, J. C. Ulwick, and R. P. Vancour</i>	1293
Air Density Variations in the Mesosphere, and the Winter Anomaly in Ionospheric Absorption <i>J. Mawdsley</i>	1298
Determination of Cloud Altitude from a Satellite..... <i>Rudolf A. Hanel</i>	1300
Observed Magnetic Declinations in West Antarctica.... <i>Ned A. Ostenso and Charles R. Bentley</i>	1301
Photomosaic Integration of Radar Precipitation Areas..... <i>F. L. Ludwig</i>	1303
Discussion of Paper by D. M. Hershfield and M. A. Kohler, 'An Empirical Appraisal of the Gumbel Extreme-Value Procedure'..... <i>Leo R. Beard</i>	1306
Authors' Reply to Discussions by I. I. Gringorten, M. A. Benson, and L. R. Beard of the Paper, 'An Empirical Appraisal of the Gumbel Extreme-Value Procedure' <i>D. M. Hershfield and M. A. Kohler</i>	1308
Discussion of Paper by P. P. Rowe, 'An Equation for Estimating Transmissibility and Coefficient of Storage from River-Level Fluctuations'..... <i>Mahdi S. Hantush</i>	1310
API Nomogram..... <i>Wayne J. Kammerer</i>	1312
Seasonal Variations in Atmospheric Carbon Dioxide Concentration..... <i>Eugene M. Wilkins</i>	1314
Addendum; The Density and Mass Distribution of Meteoritic Bodies in the Neighborhood of the Earth's Orbit..... <i>Harrison Brown</i>	1316
Note on 'The Economical Net Radiometer'..... <i>C. B. Tanner, J. A. Businger, and P. M. Kuhn</i>	1318
Comment on Paper by D. G. Singleton, 'The Geomorphology of Spread F'..... <i>John P. Miller</i>	1318
Erratum..... <i>R. A. Duncan</i>	1319
Erratum..... <i>C. B. Tanner and W. L. Pelton</i>	1319
Erratum..... <i>Mahdi S. Hantush</i>	1319

Contents

	PAGE
The Time Variations of Solar Cosmic Rays during July 1959 at Minneapolis <i>J. R. Winckler, P. D. Bhavsar, and L. Peterson</i>	99
The High-Energy Cosmic-Ray Flare of May 4, 1960. 1. High-Altitude Ionization and Counter Measurements..... <i>J. R. Winckler, A. J. Masley, and T. C. May</i>	102
The High-Energy Cosmic-Ray Flare of May 4, 1960. 2. Emulsion Measurements <i>S. Biswas and P. S. Freier</i>	102
Large-Scale Electron Bombardment of the Atmosphere at the Sudden Commencement of a Geomagnetic Storm..... <i>R. R. Brown, T. R. Hartz, B. Landmark, H. Leinbach, and J. Ortner</i>	103
The Steady State of the Chapman-Ferraro Problem in Two Dimensions..... <i>J. W. Dungey</i>	104
Radiation from a Current Filament above a Homogeneous Earth, with Application to Micropulsations <i>P. F. Law and B. M. Fannin</i>	104
Ionospheric Electron Content and Its Variations Deduced from Satellite Observations <i>K. C. Yeh and G. W. Swenson, Jr.</i>	106
A Note on 106.1-Mc Auroral Echoes Detected at Stanford Following the Solar Event of November 12, 1960..... <i>R. L. Leadabrand, W. E. Jaye, and R. B. Dyce</i>	106
A Local Reduction of <i>F</i> -Region Ionization Due to Missile Transit..... <i>Henry G. Booker</i>	107
On the Nature of Equatorial Spread <i>F</i> <i>Robert Cohen and Kenneth L. Bowles</i>	108
Doppler Shifts and Faraday Rotation of Radio Signals in a Time-Varying, Inhomogeneous Ionosphere. Part II. Two-Signal Case..... <i>John M. Kelso</i>	110
Ionization Loss Rates Below 90 Km..... <i>Cullen M. Crain</i>	111
The Relationship of Low-Height Ionosonde Echoes to Auroral-Zone Absorption and VHF <i>D</i> Scatter <i>Jens K. Olesen and J. W. Wright</i>	111
Photogrammetric Refraction Angle; Satellite Viewed from Earth..... <i>B. L. Jones</i>	111
The Diurnal Variation of <i>K</i> Indices of Geomagnetic Activity on Quiet Days in 1940-1948 <i>Seth B. Nicholson and Oliver R. Wulf</i>	111
Energy Transformation and Vertical Flux Processes over the Northern Hemisphere <i>Clayton E. Jensen</i>	111
An Experiment in the Study of the Structure of an Atmospheric System by Numerical Process <i>Stanley E. Asplund</i>	111
On the Tensor Form of Dispersion in Porous Media..... <i>Jacob Bear</i>	111
Capillary Pressure and Surface Discontinuity in Porous Media <i>Rao Channapragada and Walter Rose</i>	111
Estimating Ground-Water Recharge from Stream Hydrographs..... <i>P. Meyboom</i>	12
Interrelationships of Watershed Characteristics..... <i>Don M. Gray</i>	12

(Continued inside back cover)